https://doi.org/10.1093/plphys/kiac180 PLANT PHYSIOLOGY 2022: 189: 1625–1638

Plant Physiology®

Research Article

Single-parent expression complementation contributes to phenotypic heterosis in maize hybrids

Jutta A. Baldauf [1], Meiling Liu [2,‡], Lucia Vedder [3], Peng Yu [4], Hans-Peter Piepho [5], Heiko Schoof [3], Dan Nettleton [2] and Frank Hochholdinger [1,*,†]

1 Institute of Crop Science and Resource Conservation, Crop Functional Genomics, University of Bonn, 53113 Bonn, Germany
2 Department of Statistics, Iowa State University, Ames, Iowa 50011-1210, USA
3 Institute of Crop Science and Resource Conservation, Crop Bioinformatics, University of Bonn, 53115 Bonn, Germany
4 Emmy Noether Group Root Functional Biology, Institute of Crop Science and Resource Conservation, University of Bonn, 53113 Bonn, Germany
5 Institute of Crop Science, Biostatistics Unit, University of Hohenheim, 70599 Stuttgart, Germany

*Author for communication: hochholdinger@uni-bonn.de
†Senior author.
‡Present address: Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, 98109, USA.

J.A.B. carried out the experiments, conducted statistical analyses, interpreted the data, and drafted the article. M.L. and D.N. performed the gene activity analyses. L.V. and H.S. were involved in the bioinformatics handling of the data. H.P.P. provided the experimental design for the RNA-seq experiment and helped with statistical analyses. F.H. conceived and coordinated the study, participated in data interpretation, and drafting the article. All authors edited the article and approved the final draft.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plphys/pages/general-instructions) is: Frank Hochholdinger (hochholdinger@uni-bonn.de).

Abstract

The dominance model of heterosis explains the superior performance of F_1-hybrids via the complementation of deleterious alleles by beneficial alleles in many genes. Genes active in one parent but inactive in the second lead to single-parent expression (SPE) complementation in maize (*Zea mays* L.) hybrids. In this study, SPE complementation resulted in approximately 700 additionally active genes in different tissues of genetically diverse maize hybrids on average. We established that the number of SPE genes is significantly associated with mid-parent heterosis (MPH) for all surveyed phenotypic traits. In addition, we highlighted that maternally (SPE_B) and paternally (SPE_X) active SPE genes enriched in gene co-expression modules are highly correlated within each SPE type but separated between these two SPE types. While SPE_B-enriched co-expression modules are positively correlated with phenotypic traits, SPE_X-enriched modules displayed a negative correlation. Gene ontology term enrichment analyses indicated that SPE_B patterns are associated with growth and development, whereas SPE_X patterns are enriched in defense and stress response. In summary, these results link the degree of phenotypic MPH to the prevalence of gene expression complementation observed by SPE, supporting the notion that hybrids benefit from SPE complementation via its role in coordinating maize development in fluctuating environments.

Introduction

Maize (*Zea mays* L.) is naturally an outcrossing species, which resulted in an exceptional degree of intraspecific structural genomic diversity during domestication (Springer et al., 2009; Swanson-Wagner et al., 2010; Jiao et al., 2017). The maize genome contains two evolutionarily distinct subgroups of genes classified by the presence or absence of syntenic orthologs in the closely related sorghum (*Sorghum*

Received December 09, 2021. Accepted March 23, 2022. Advance access publication April 21, 2022

Open Access

bicolor L.) genome (Schnable and Lyons, 2011). Evolutionarily older syntenic genes can be identified by the presence of syntenic orthologs in sorghum (Schnable and Lyons, 2011). In contrast, evolutionarily younger non-syntenic genes lack syntenic orthologs in sorghum as they evolved by gene duplications after the divergence of the maize and sorghum lineages (Schnable and Lyons, 2011).

Highly heterozygous F_1-hybrids are more vigorous than their homozygous, genetically distinct parental inbred lines—a phenomenon known as heterosis (Shull, 1948). Genetic diversity between the parental inbred lines is essential to fully exploit the effect of hybrid vigor. The effect of heterosis is most evident for traits at the adult stage of plant development, like crop yield, fertility, and biomass (Flint-Garcia et al., 2009). Total grain yield in maize is mainly determined by the highly heterotic traits ear length, ear diameter, number of kernels per ear, and thousand-kernel weight (Austin and Lee, 1998; Wang et al., 2007; Flint-Garcia et al., 2009; Zhang et al., 2021).

Variation in transcriptional regulation is a major principle of phenotypic diversity and a potential source of heterosis (Botet and Keurentjes, 2020).

In maize primary roots, hundreds of genes are expressed in hybrids but in only one of its two parental inbred lines (Paschold et al., 2012, 2014; Marcon et al., 2017; Baldauf et al., 2018). This specific type of gene expression complementation was denoted as single-parent expression (SPE; Paschold et al., 2012). SPE describes an extreme instance of gene expression complementation in maize hybrids and is consistent with the classical dominance model of heterosis. The dominance model, also known as the complementation model, describes the concept that deleterious alleles of different genes in both inbred lines are complemented in the F_1-hybrid by beneficial alleles, which results in its superior phenotypic performance (Jones, 1917). Recently, SPE complementation was demonstrated in genetically diverse maize genotypes along different stages of primary root development grown under controlled environmental conditions (Baldauf et al., 2018) and in aboveground tissues (Li et al., 2021). It was shown that non-syntenic genes were the driving force of gene expression complementation in hybrids (Paschold et al., 2014; Marcon et al., 2017; Baldauf et al., 2018; Li et al., 2021).

The objective of this study was to examine the role of SPE complementation in the phenotypic manifestation of heterosis in a genetically diverse panel of genotypes to survey how gene co-expression networks are connected with SPE and phenotypic traits.

Results

Data quality and gene activity status

In this study, we examined the transcriptomic plasticity of six maize F_1-hybrids and their parental inbred lines in three aboveground organs. We selected six paternal inbred lines (A554, H84, H99, Mo17, Oh43, and W64A) from different clades of the phylogenetic tree of maize (Lorenz and Hoegemeyer, 2013; see Figure 1a in Baldauf et al., 2018) and used them as male pollen donors in crosses with the inbred line B73 as a common female. In this study, the closely related inbred lines B73 and H84 belong to the stiff stalk germplasm pool, while the five other genotypes were selected from the non-stiff stalk germplasm pool. The inbred line B73 was selected as a common female parent because this vigorous genotype is the female parent of choice in many hybrid breeding programs and the reference-quality genome sequence of this genotype is available (Jiao et al., 2017). The diverse maize genotypes were grown according to a randomized block design in a field experiment (Supplemental Figure S1A). To capture the transcriptomic dynamics in vegetative and generative organs, the youngest fully developed leaf, that is, the flag leaf, the meiotic tassel, and the immature ear before pollination (Figure 1A), were harvested from each genotype. For each genotype-by-organ combination four biological replicates were subjected to RNA-sequencing (RNA-seq). However, for the common maternal inbred line B73, we surveyed 12 biological replicates per organ to optimize the precision of direct comparisons between B73 and the remaining genotypes by decreasing the standard error.

To check the quality and the sample relations in the dataset, we performed a principal component analysis (PCA) of all active genes across all genotypes and organs (Figure 1B; Supplemental Figure S2). Furthermore, we included a previously published primary root dataset of the same genotype panel (Baldauf et al., 2018) to display transcriptomic relationships from early vegetative belowground to aboveground vegetative and generative organs. The biological replicates of the different genotypes form distinct organ-specific clusters, indicating that the transcriptomes of the different genotypes have a higher transcriptomic similarity than the different organs of the same genotype (Supplemental Figure S2). The first three principal components (PCs) explain in total 60% of the variance in both datasets. The first PC separates the belowground primary root from the ear, tassel, and leaf transcriptome, while PC 2 and PC 3 separate the two aboveground generative organs (i.e. ear and tassel) from the two vegetative organs (i.e. root and leaf) (Supplemental Figure S2). Between both datasets, only 5% of the variance (PC 4) can be attributed to the transcriptomic distinctiveness between the different genotypes (Figure 1B). For each organ, the samples of the maternal inbred line B73 form a distinct cluster and are separated from the distantly related paternal inbred lines (Figure 1B). The sample clusters of the hybrids are located between the clusters of both their parental inbred lines. Hence, this reflects the genotypic relatedness among the diverse panel of genotypes and highlights the high quality of the data.

Aboveground tissues display SPE complementation patterns enriched in non-syntenic genes

Genes that are active in one parental inbred line and the hybrid, but inactive in the second parental inbred line, display

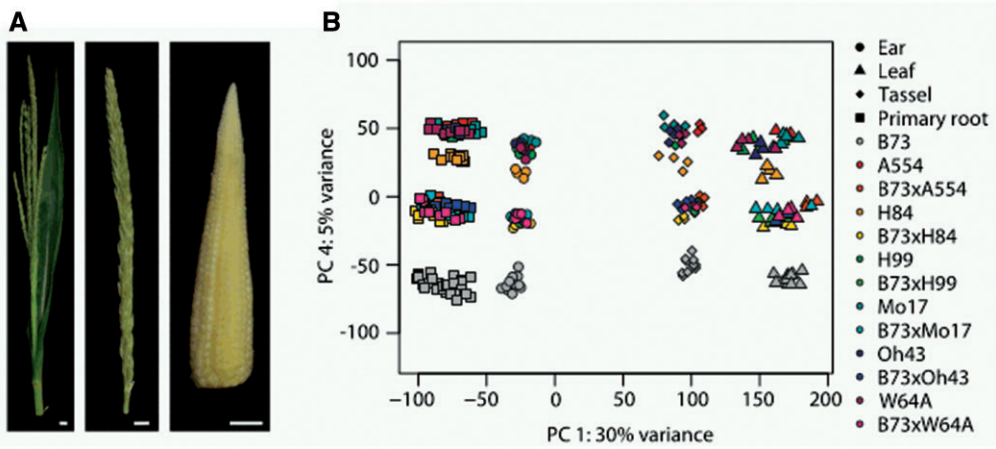

Figure 1 Plant material surveyed in this study and RNA-seq sample relationship. A, The three aboveground organs used for RNA-seq: flag leaf (youngest fully developed leaf), main axis of the meiotic tassel and unpollinated ear. Scale = 1 cm. B, PC plot of the first and fourth component of all expressed genes in all 180 RNA-seq samples of the three aboveground organs and of a previously published RNA-seq dataset of maize primary roots (Baldauf et al., 2018). Each genotype is highlighted by a different color. Hybrids are represented in lighter shade of the same color as the paternal inbred line. The four different tissues are distinguished by the shapes of their features.

SPE. Such SPE patterns are an extreme instance of gene expression complementation. We distinguish between two SPE patterns: SPE_B genes are active in the common maternal inbred line B73 but inactive in the corresponding paternal inbred line. Vice versa, SPE_X genes are active in the hybrid and the paternal inbred line X (*X* = A554, H84, H99, Mo17, Oh43, or W64A) but not in the maternal inbred line B73. For each aboveground organ, we determined the number of genes displaying SPE_B and SPE_X patterns for all six parent–hybrid triplets (Figure 2A). Among all genotype combinations, we observed between 1,196 and 2,026 SPE patterns. On average, fewer genes displayed SPE_B expression patterns across all genotype-by-organ combinations in comparison to the number of genes displaying SPE_X patterns. In the three different aboveground organs, on average, we identified the highest number of SPE genes in the tassel, followed by the ear and by the flag leaf (Figure 2A).

The genome of modern maize contains two subsets of genes: evolutionarily older syntenic genes and evolutionarily younger non-syntenic genes. In total, 56% of the annotated maize genes (B73v4) are syntenic genes, while 44% are non-syntenic. In all genotype-by-organ combinations, the evolutionarily younger nonsyntenic genes were significantly enriched among SPE genes (56%–84%, Figure 2B, colored bars) in comparison to their prevalence among all active genes (8%–11%; Figure 2B, gray bars) in the corresponding aboveground organ (Figure 2B).

Hybrids express more genes than their parental inbred lines

Within each parent–hybrid triplet, the hybrids express more genes than their parental inbred lines in all three examined aboveground shoot organs (Figure 2C). The only exception is the hybrid B73xH99, which exhibits fewer active genes than its paternal inbred line H99 in the flag leaf. On average, hybrids express 957 genes more than their parental inbred lines in ears, 424 (457 excluding H99) more genes in flag leaves, and 829 more genes in tassels (Figure 2C).

Among the different organs, the highest numbers of active genes were observed in both generative organs, the meiotic tassel and the unpollinated immature ear followed by the flag leaf samples.

SPE genes are highly tissue- and genotype-specific

Across all six parent–hybrid triplets, the genes displaying SPE_B and SPE_X patterns were highly tissue-specific (Figure 3A). In contrast, a smaller proportion of genes displayed conserved SPE_B and SPE_X patterns among the three tissues (Figure 3A). Among all expressed genes determined in each of the 13 genotypes, between 81% and 90% were expressed in all three aboveground tissues, while only between 2% and 7% were tissue- specific (Supplemental Figure S3).

Similarly, genes displaying SPE patterns displayed considerable parent–hybrid triplet specificity and a lower conservation between the triplets (Figure 3B). In general, SPE_B genes displayed a higher genotype specificity than their SPE_X counterparts, because SPE patterns of genes only present in X were not captured.

The number of SPE genes is significantly correlated with MPH

During the harvest of the plant material for the RNA-seq experiment, we recorded several phenotypic traits including the length of the unpollinated immature ear, meiotic terminate tassel length, flag leaf length, plant height, and the number of leaves and stem nodes. At maturity of the remaining plants in each plot, we harvested the open-pollinated mature ears and determined their length. For each trait, the level of mid-parent heterosis (MPH) was estimated in each parent–hybrid triplet (Supplemental Figure S4; Supplemental Table S1). On average, the size of the

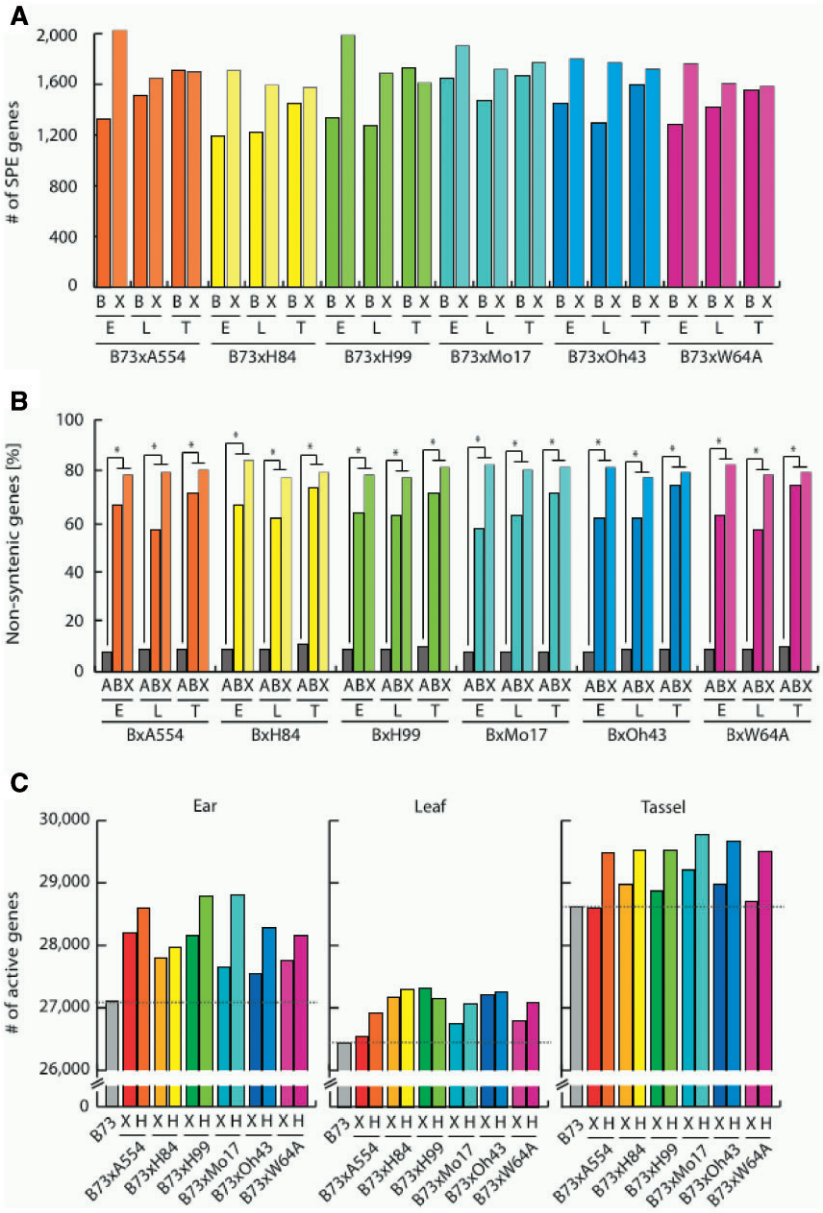

Figure 2 SPE complementation in three aboveground tissues. A, Number of SPE_B (B) and SPE_X (X) genes in six parent–hybrid triplets. E, L, and T indicate the analyzed aboveground tissues: E = unpollinated ear, L = flag leaf, T = meiotic tassel. B, Proportion of nonsyntenic genes among SPE_B (B) and SPE_X (X) patterns and among the total number of active genes that are not SPE (A, gray) per parent–hybrid triplet for each of the three analyzed aboveground tissues (E, L, and T). Significant differences (Fisher's exact test $\alpha = 0.05$) are indicated by an asterisk. C, Number of active genes in hybrids (H) and their parental inbred lines in each aboveground tissue. For each parent–hybrid comparison, the corresponding paternal inbred line is abbreviated by "X" (X = A554, H84, H99, Mo17, Oh43, and W64A). The thin dotted line indicates the number of active genes in the inbred line B73.

mature ear (39%) and plant height (34%) displayed the highest levels of MPH across all hybrids. In contrast, the number of leaves (7%) and flag leaf size (17%) displayed lower levels of MPH. Irrespective of the trait, the hybrid B73xH84 displayed on average the lowest level of MPH (15%) reflecting the close relatedness of its parental inbred lines (Lorenz and Hoegemeyer, 2013; Supplemental Figure S4).

A regression analysis of the level of MPH for the different phenotypic traits on the number of genes displaying either the SPE_B or SPE_X pattern in the aboveground part of the plant revealed in all instances a significant positive association (Figure 4; Supplemental Table S2). Plant height and ear size displayed the highest levels of MPH (Supplemental Figure S4).

In addition, we performed a regression analysis of previously identified SPE genes in primary maize roots in the same genotype panel (Baldauf et al., 2018) to the phenotypic aboveground organs (Figure 4B; Supplemental Table

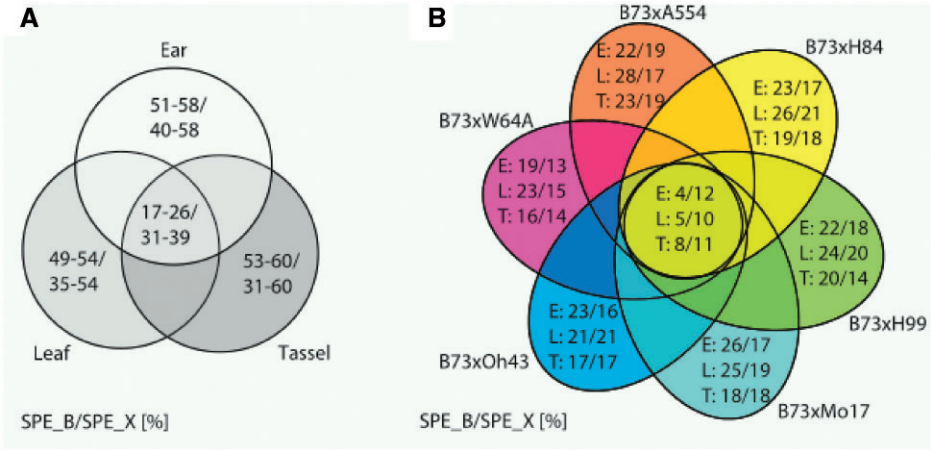

Figure 3 Tissue- and genotype-specific SPE patterns in three aboveground tissues. The Venn diagram in (A) displays the proportion of SPE_B and SPE_X genes, which are tissue-specific and conserved between the three aboveground tissues. The numbers on the left side of the slash indicate the ratio of the SPE_B genes among the six different parent–hybrid triplets, while the numbers on the right side of the slash indicate the ratio of the SPE_X genes. In (B), the genotype-specific SPE_B and SPE_X patterns and their overlap between the six different hybrids is displayed for the unpollinated ear (E), the flag leaf (L), and meiotic tassel (T). On the left side of the slash, the percentage of SPE_B genes is indicated and on the right side, the percentage of the SPE_X genes.

S2) to test if SPE genes determined early in seedling development can be used to predict MPH at the adult stage of the plant. Similar to the results of the regression of the SPE genes determined in aboveground organs at the adult stage, the SPE genes determined in primary roots were significantly associated with the level of MPH for all examined traits (Supplemental Table S2).

SPE_B and SPE_X patterns enriched in gene co-expression modules are highly correlated within each SPE pattern but clearly separated between these SPE patterns

For each of the three aboveground organs, we conducted a weighted gene co-expression network analysis (WGCNA) to identify clusters of highly connected genes, which are related to expression patterns and phenotypic traits determined in this study. Using dynamic tree cut, we grouped the hierarchically clustered expression profiles into individual modules representing clusters of highly correlated genes. To each module, we assigned a specific color name to distinguish between different modules. For each organ, we constructed the co-expression network for only the highly variable expressed genes among the prefiltered active genes including 6,232 genes for the unpollinated immature ear, 5,916 genes for the flag leaf, and 6,570 genes for the meiotic tassel. Setting the minimum module size to 30 genes per module, we identified 51, 55, and 52 modules in the ear, leaf, and tassel dataset, respectively. Among the different organs, up to 627 (ear), 591 (leaf), and 1,057 (tassel) genes were assigned to the largest module (= turquoise module) (Supplemental Data set 1).

The module eigengene is defined as the first PC of a given module and, thus, serves as a representative of the gene expression profile in that given module. Correlating the module eigengenes to the sample information represented by the family structure in each organ revealed several modules, which display high correlation values in the four biological replicates of a specific paternal inbred or maternal line and its hybrid, but lower correlation values in the remaining genotypes including the second parental inbred line (Supplemental Figure S5). Examining the expression patterns of the genes within each module, we identified SPE_B-enriched and SPE_X-enriched modules (SPE pattern determined in five or more of the six hybrids) (Supplemental Table S3) in each organ. In each organ, SPE_B-enriched and SPE_X-enriched modules are clearly separated (Figure 5, A–C). This indicates that the gene expression profiles of the SPE_B and SPE_X genes are highly correlated but distinct from each other.

Besides the SPE_B-enriched and SPE_X-enriched modules, we also identified 12 modules in which a specific SPE pattern, identified in a single parent–hybrid triplet, is enriched (Supplemental Table S3).

Highly connected SPE genes display significant phenotypic trait correlations

To close the gap between the phenotypic superiority of hybrids in the field and their gene expression patterns, we correlated the clusters of highly connected genes, represented by their expression profiles, with phenotypic traits. To this end, we identified several modules, which displayed significant associations (correlation value $\geqslant 0.5$, $P \leqslant 0.001$) to phenotypic traits. The resulting significant module–trait associations are depicted in Figure 5, D–F for the SPE_B- and SPE_X-enriched modules and in Supplemental Figure S6 for all identified modules in each dataset. Interestingly, the SPE_B- and SPE_X-enriched modules, which are significantly associated with a specific trait, show correlation values in opposite directions. Overall, the SPE_B-enriched modules showed a positive correlation in almost all instances, while

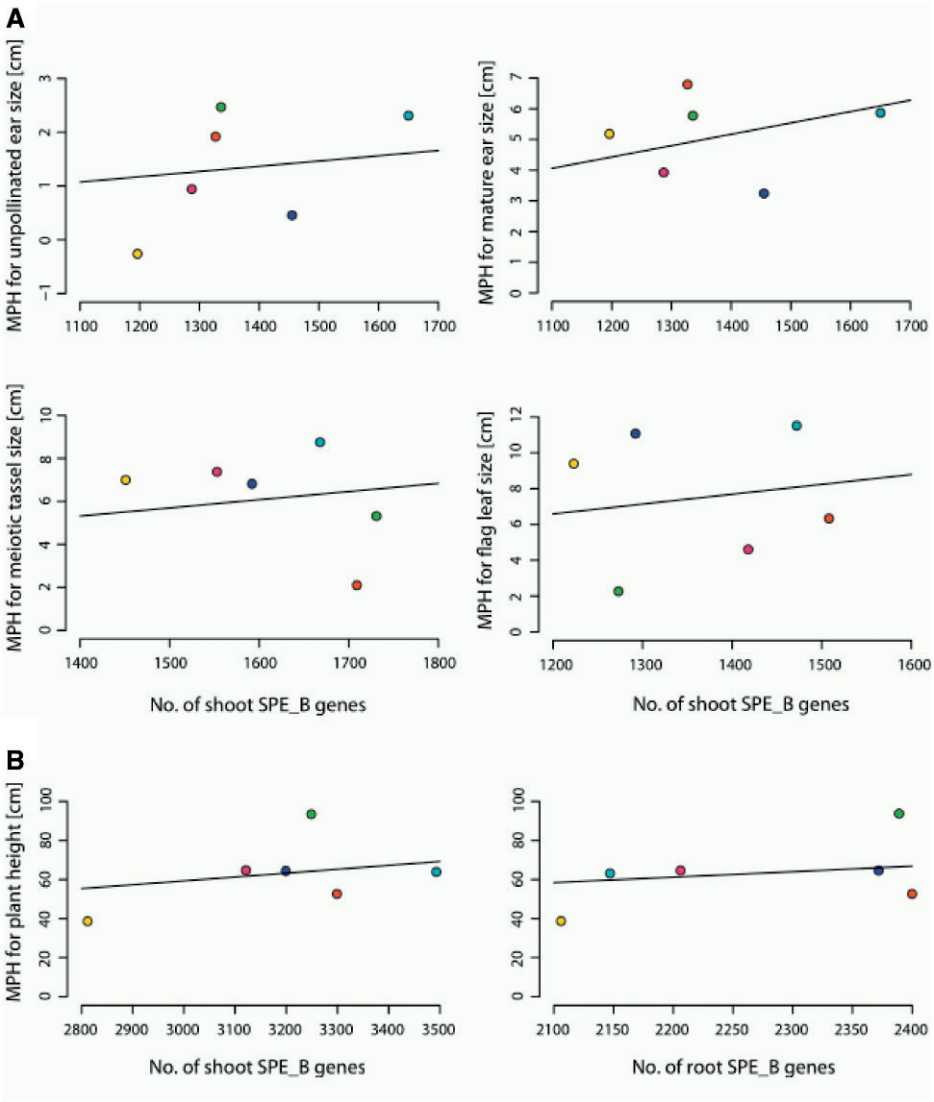

Figure 4 SPE genes are highly correlated with MPH. In (A), the results of the regression analysis of the number of SPE_B genes, identified in the corresponding aboveground shoot datasets, on MPH are displayed for the size of the following traits: unpollinated ear, mature ear, meiotic tassel, and flag leaf. The left part of (B) displays the regression line of the number of SPE_B genes identified in the aboveground tissue dataset on the MPH of the trait plant height. In the right part of B, the regression line of the SPE_B genes identified in the previously published primary root dataset (Baldauf et al., 2018) on plant height is displayed. For each hybrid, the number of SPE_B genes and the corresponding level of MPH for a specific trait are indicated by color-coded data points: red = B73 × A554, yellow = B73 × H84, green = B73 × H99, turquoise = B73 × Mo17, blue = B73 × Oh43, magenta = B73 × W64A.

the SPE_X-enriched were in general negatively correlated with the phenotypic traits. In addition, the modules enriched by a specific SPE pattern show in most cases a significant correlation to individual traits and contrasting correlation values between SPE_B and SPE_X patterns of the same genotype (Supplemental Figure S6).

Functional characterization of modules displaying significant trait correlations

We performed a gene ontology (GO) enrichment analysis to identify overrepresented biological and molecular functions among the genes assigned to the SPE-enriched modules. For some modules, we identified several significantly enriched GO terms (Supplemental Table S4). Among the SPE_X-enriched modules identified in all three organs (brown in ear, red in leaf, brown in tassel), the GO term "regulation of ATPase activity" was consistently identified. In addition, GO terms related to the biological processes "response to heat," "response to high light intensity," and "defense response signaling pathway, resistance gene-dependent" (a synonym for GO term GO:0002758 "innate immune response-activating signal transduction") and diverse subcategories were enriched among these modules (Supplemental Table S4). Furthermore, we identified especially among the SPE_B-enriched modules in the leaf and tassel dataset, GO

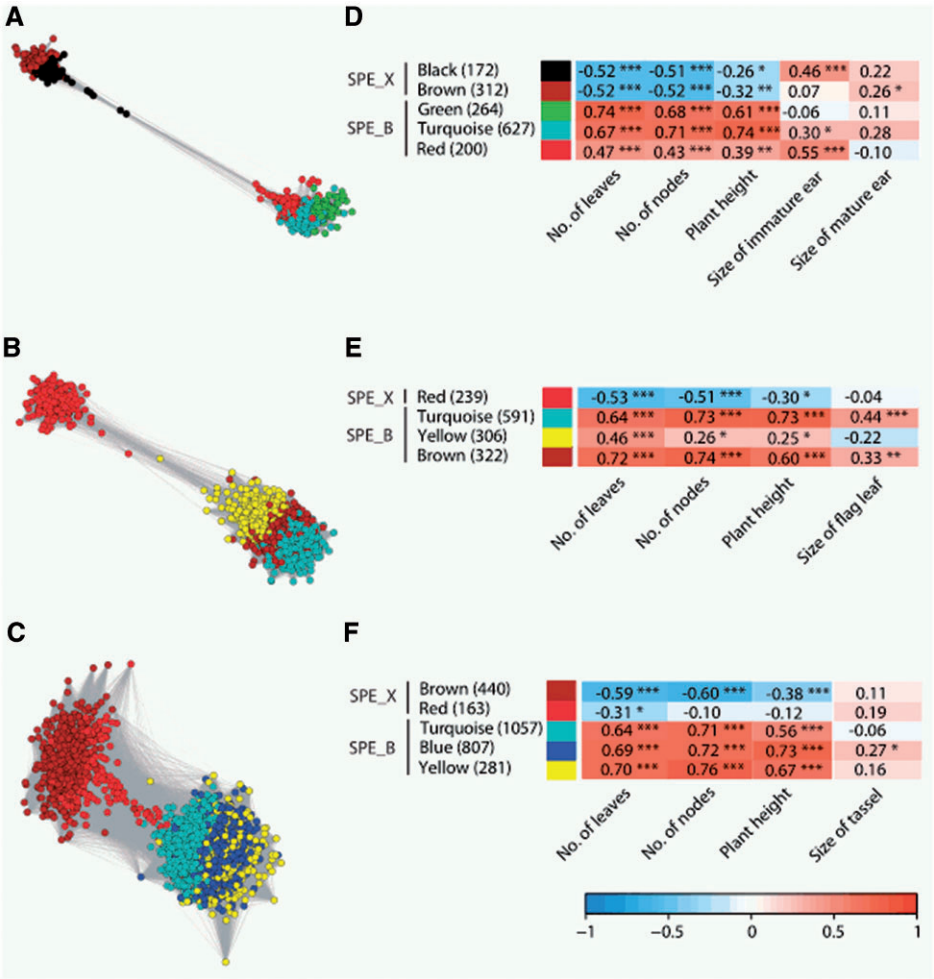

Figure 5 Modules enriched by SPE patterns and their trait associations. In the left part of the figure (A–C), the relation of the modules generally enriched in either SPE_B or SPE_X genes identified in five or more of the six parent–hybrid triplets is displayed in A the unpollinated ear, (B) the flag leaf, and (C) the meiotic tassel. The heatmaps in the right part of the figure (D–F) display correlation values of the module eigengenes to phenotypic traits identified in (D) the unpollinated ear, (E) the flag leaf, and (F) the meiotic tassel. Only modules which are generally enriched in genes displaying either SPE_B or SPE_X patterns identified in five or more of the six parent–hybrid triplets are shown. The results for all identified modules are displayed in Supplemental Figure S6. Numbers in brackets indicate the number of genes in each module. The heatmaps display Pearson correlation values and asterisks indicate significant levels: *α = 0.05, **α = 0.01, ***α = 0.001.

terms associated to "cell wall organizations and biogenesis," "pollen tube development," "cell development," and "growth" (Supplemental Table S4).

Discussion

Maize exhibits an exceptional degree of intra-specific genomic diversity (Swanson-Wagner et al., 2010; Marroni et al., 2014) that relates to transcriptomic diversity (Hansey et al., 2012; Hirsch et al., 2014; Jin et al., 2016). In this study, we applied RNA-seq to survey how the genomic diversity of maize is reflected in the transcriptomic landscape of F_1-hybrids and their parental inbred lines in mature, vegetative, and generative tissues (Figure 1A).

To reflect a maximum genetic diversity in the hybrids, we crossed the inbred line B73 with five paternal inbred lines, representing different clades of a published phylogenetic tree and with the inbred line H84 of the same clade (Lorenz and Hoegemeyer, 2013). The result of our PCA confirmed the expected kinship relations (Figure 1B; Supplemental Figure S2). Hybrids share 50% of their genomes with each of their parental inbred lines. Therefore, the hybrid samples were located between the parental samples. Consistent with the pedigree lineage of the surveyed inbred lines (Troyer, 2001), the stiff stalk line H84 displayed the closest relation to B73, while the non-stiff stalk lines were more distantly related to B73. Similar clustering was previously observed in primary roots of maize (Baldauf et al., 2018). Although the different genotypes were transcriptomically distinct, a much higher proportion of the variance in our dataset is attributed to the dissimilarity in the transcriptomes among the examined tissues (Figure 1B; Supplemental S2). A similar transcriptomic separation of five different maize tissues was observed in a study that surveyed five maize tissues in a large set of genotypes (Li et al., 2021). In contrast, different

stages of primary root development were not transcriptomically separated as observed for distinct tissues (Baldauf et al., 2018; Figure 1B).

The concept of SPE complementation implies that genes active in hybrids are only transcribed in one of the parental inbred lines (Paschold et al., 2012). SPE patterns were observed throughout seedling root (Paschold et al., 2012, 2014; Baldauf et al., 2018) and mature vegetative and generative aboveground development under field conditions (Figure 2A; Baldauf et al., 2018; Li et al., 2021).

We demonstrated that gene expression complementation in F_1-hybrids is a general mechanism to increase the total number of active genes in maize hybrids compared to their parental inbred lines (Figure 2C). In maize primary roots, the total number of expressed genes in the hybrids and their parental inbred lines increased consistently across all genotypes along development (Baldauf et al., 2018), reflecting the increase of morphological complexity in the examined roots (Hochholdinger et al., 2004). In this study of the transcriptomic plasticity in aboveground tissues, we observed a similar pattern: we identified the highest number of active genes in the meiotic tassel followed by the immature ear. We observed less expressed genes in the fully developed flag leaf (Figure 2C). The observed range is consistent with previous RNA-seq studies in B73 (Davidson et al., 2011; Stelpflug et al., 2016). The highest number of expressed genes in the meiotic tassel, which was harvested prior to pollen shedding, can be attributed to the unique expression profile of pollen-specific genes as previously demonstrated (Davidson et al., 2011). Hence, the observed expression profiles reflect the transition from vegetative to generative development.

The evolutionary age of maize genes can be estimated based on their synteny to genes in the unduplicated, closely related sorghum genome (Schnable and Lyons, 2011). Maize genes, which are syntenic to sorghum genes, are evolutionarily older than non-syntenic maize genes that have no sorghum homolog. In B73v4, ~54% of protein-encoding genes are syntenic, while ~44% are non-syntenic (Schnable, 2019). It was suggested that non-syntenic genes often contribute to the capacity of plants to adapt to fluctuating environmental conditions (Schnable, 2015). Among all genes expressed in the parent–hybrid triplets, between 8% and 11% are non-syntenic and thus underrepresented compared to their genomic prevalence (Figure 2B). Similar to previous observations in maize (Paschold et al., 2014; Baldauf et al., 2018; Li et al., 2021), the identified SPE patterns in this study were significantly enriched in non-syntenic genes (56%–84%) in each parent–hybrid triplet (Figure 2B). Since maize hybrids can cope better with environmental fluctuations than their parental inbred lines (Betrán et al., 2003), these results support the notion that an overrepresentation of non-syntenic SPE genes could contribute to the adaptability of hybrids. This hypothesis is further underscored by the observation that SPE genes are very stable in their expression pattern under water deficit conditions (Marcon et al., 2017).

Heterosis can be monitored throughout plant development, but the degree of the phenotypic advantage is highly trait- and genotype-specific (Hochholdinger and Baldauf, 2018). In general, continuous traits (e.g. plant height or yield) display higher levels of heterosis than discrete traits (e.g. number of leaves or stem nodes) (Mehta and Sarkar, 1992; Flint-Garcia et al., 2009; Yi et al., 2019). Both plant height and grain yield, as examples of continuous traits, are significantly correlated and display among the highest levels of heterosis (Yang et al., 2017). Among the traits examined in this study, plant height and size of the mature ear displayed the highest level of heterosis (Supplemental Figure S4). The concept of heterotic groups (Melchinger and Gumber, 1998) is based on the observation that increasing the genetic diversity of the parental inbred lines via their selection from two different heterotic pools leads to superior performance of their hybrid progeny (Larièpe et al., 2012). In our study, crosses between the stiff stalk line B73 and the five non-stiff stalk lines from other clades (Lorenz and Hoegemeyer, 2013) displayed on average the highest level of phenotypic heterosis (Supplemental Figure S4). In contrast, the hybrid B73 × H84 with two stiff stalk parents displayed on average the lowest levels of MPH (Supplemental Figure S4). Likewise, the number of SPE genes varies highly between the different hybrids, but still reflects the relatedness of the parental inbred lines (Figure 2A). A correlation between genetic distance and the number of SPE genes has also been suggested by Li et al. (2021). Hence, SPE complementation might be one of the links that converts the genetic distance of the parents into phenotypic heterosis in hybrids as demonstrated in this study. Consistent with the variation of the level of MPH between different genotypes and traits, SPE genes are also highly genotype- and tissue-specific (Figure 3). Nevertheless, the number of SPE genes is significantly positively correlated with the level of MPH irrespective of the function of these genes (Figure 4; Supplemental Table S2). These results link to our observation that SPE genes are significantly enriched with non-syntenic genes, which could contribute to the superior performance of the hybrids at fluctuating environmental conditions. Overall, in accordance with the classical dominance model (Jones, 1917), gene expression complementation in hybrids is associated with heterosis.

Interestingly, SPE genes determined in belowground primary root tissues are also significantly correlated with MPH observed for mature, vegetative, and generative traits (Figure 4B; Supplemental Table S2). Although further studies are necessary to examine the role of SPE genes in heterosis manifestation, these results might provide an interesting tool to predict MPH based on SPE genes identified during early seedling root development. Due to their simple structure, maize seedling roots are an attractive model for studying heterosis under controlled environmental conditions in high throughput (Hoecker et al., 2006; Paschold et al., 2010).

Although we have seen that SPE genes are significantly correlated with the level of MPH for all examined traits, we

observed that specific groups of SPE genes have a stronger effect on the phenotype of the plants (Figure 5, D–F; Supplemental Figure S6). In the gene co-expression analysis, SPE_X-enriched modules were in most cases negatively correlated with the observed phenotypic traits while SPE_B-enriched modules showed positive correlations. This pattern is also reflected in the regression coefficients (Supplemental Table S2) and might reflect either the functional contribution of SPE_B and SPE_X genes to the observed phenotypes or the parental contribution of individual SPE patterns to their hybrid offspring. The yellow (in leaf) and the turquoise (in tassel) modules, in which SPE_B genes are overrepresented in five or more of the six parent–hybrid triplets, are predominantly enriched in growth-promoting GO terms (Supplemental Table S4). This indicates that the genes assigned to these modules might promote the fitness of these plants, which is also reflected in a significantly positive association of the module eigengenes of these SPE_B-enriched modules to the phenotypic traits (Figure 5, D–F). In contrast, among the SPE_X-enriched modules GO terms related to defense- and stress-related biological processes were specifically enriched, which might explain the negative correlation of the corresponding module eigengenes to the measured phenotypic traits (Supplemental Table S4; Figure 5, D–F). Although a competition between plant immunity and growth due to limited resource allocations was reported previously (Denancé et al., 2013; Huot et al., 2014), it was demonstrated that a reduction of the basal defense response in Arabidopsis (*Arabidopsis thaliana*) F_1-hybrids promoted leaf growth but did not necessarily compromise their stress tolerance (Groszmann et al., 2015; Gonzalez-Bayon et al., 2019). There seems to exist a cross-talk between the networks promoting the response of the plant to environmental stimuli and to growth. In Arabidopsis hybrids, it was shown that stress-responsive genes were repressed under normal conditions, but diurnally induced under stress conditions and thereby balancing the tradeoff between stress response and growth (Miller et al., 2015). In combination with our observation that both SPE_B and SPE_X patterns are significantly correlated with MPH, this supports the notion that hybrid plants benefit from both SPE patterns in fluctuating environmental conditions.

Furthermore, the parental contribution of individual SPE genes is, for example, reflected in the significant trait association of the dark-red and pink module identified in the dataset of the unpollinated ear. The parental inbred line H84 had on average 2 cm shorter immature ears in comparison to the maternal inbred line B73 at harvest (Supplemental Table S1). The dark-red module, which is specifically enriched in SPE_B genes determined in the parent–hybrid triplet B73xH84, is positively correlated with the size of the immature ear. In contrast, the pink module, which is enriched by SPE_X genes identified in the same hybrid, is negatively correlated with that trait (Supplemental Figure S6). This indicates that the maternal inbred line B73 has a higher contribution to the size of the immature ear of the hybrid B73 × H84 than the inbred line H84. Although estimating parental effects on the transcriptome of the hybrid is generally challenging (Botet and Keurentjes, 2020), these results are in line with the observation that the level of heterosis depends highly on the measured trait and environmental effects (Li et al., 2018).

Conclusions

In summary, our results suggest that SPE complementation contributes to phenotypic heterosis in maize hybrids and that genes displaying this type of gene expression pattern could be a promising target for maize breeding.

Materials and methods

Plant materials and field experiments

We crossed the six male maize (*Zea mays* L.) inbred lines A554, H84, H99, Mo17, Oh43, and W64A with the common female parental inbred line B73 to obtain six F_1-hybrids. The field experiment was laid out as a randomized block design with four blocks at the experimental field station Endenich of the University Bonn, Germany (50°43′47.8″N 7°04′31.1″E) in the summer of 2018 (Supplemental Figure S1A). Each experimental unit was a 3.2 m long plot consisting of four rows with a row spacing of 0.65 m. In each row we grew 15 plants with a planting distance of 0.2 m between adjacent plants. Within each block, 15 plots representing the six hybrids, six paternal inbred lines and three times the common maternal inbred line B73 were randomly arranged. For B73, we generated in total 12 biological replicates, amounting to three plots per block of the design. These additional replications accounted for the fact that B73 is involved in all 6 heterosis contrasts, whereas all other genotypes, for which we generated 4 biological replicates, are present in only one heterosis contrast (Piepho, 2005). We harvested the immature unpollinated ear, the main terminal axis of the meiotic tassel, and 10 cm of the flag leaf tip between vegetative stage 9 and 16, i.e. between the expansion of leaf no. 9 and 16, when the flag leaf released the tassel (BBCH Stage 55) and the terminating stem node, but prior to pollen shedding (Figure 1A). To prevent marginal plot effects, we harvested and pooled plant material of five randomly selected plants of the two middle rows of a plot for each biological replicate. At harvest date, we phenotyped the plants of which the RNA samples were taken for the following traits: plant height, number of leaves, number of stem nodes, and total length of harvested organs which were used for RNA-seq (immature unpollinated ear, main terminal tassel and flag leaf). In addition, open-pollinated mature ears of the remaining plants (n = 10) of each plot were harvested to determine the length of the mature ears.

Analysis of phenotypic data

For each phenotypic trait of each parent–hybrid triplet, we estimated the level of MPH based on a linear mixed model reflecting the genotype factor and block factors of the

randomized block design using the lme4 package in R (Bates et al., 2015):

$$Y_{ijk(m)} = \mu + g_m + b_i + p_j + e_{ijk(m)}$$

where $Y_{ijk(m)}$ represents the mean phenotypic value of a specific trait of interest of the respective genotype m, μ represents the intercept, g_m represents the fixed effect for genotype m, b_i represents the fixed effect for block i, p_j represents the random effect for plot j nested within block i, and $e_{ijk(m)}$ represents the random error effect for plant k of genotype m in block i and plot j. We calculated the levels of MPH based on the estimated least-square means of each genotype as the difference of the performance of the hybrid (F_1) and the average performance of its two parental inbred lines ((P_1 + P_2)/2) (Hochholdinger and Hoecker, 2007).

RNA isolation and experimental design of RNA-seq

We ground plant material in liquid nitrogen and isolated total RNA with the RNeasy Plant Mini kit (Qiagen, Venlo, The Netherlands). For all samples, we detected a RIN (RNA integrity number, Schroeder et al., 2006) value $\geqslant$8.3 with a Bioanalyzer (Agilent RNA 6000 Nano Chip; Agilent Technologies, Santa Clara, CA, USA). cDNA libraries were constructed with the TruSeq RNA sample preparation kit (Illumina, San Diego CA, USA). For sequencing, each library was indexed with a distinct Illumina TruSeq adapter.

We used the randomized block design of the field experiment to optimize the RNA-seq experiment represented by a row–column design, in which the rows represent the lanes of the flow cells and the columns represent the blocks (Supplemental Figure S1B). Samples harvested from the same field block remained in the same block for the RNA-seq experiment. This helped to account for repeated measurements on the same experimental unit and to obtain a higher precision for all pairwise comparisons involving the three genotypes from the same cross (parents and hybrid) in subsequent statistical analyses. The combination of a hybrid with its male parent is denoted as a "pair". In total, six pairs (A554 and B73xA554, H84 and B73xH84, H99 and B73xH99, Mo17 and B73xMo17, Oh43 and B73xOh43, W64A and B73xW64A) were analyzed. To each lane (≙ row), we assigned two pairs of two different organs per field block, while the pairs of the two organs were identical (16 libraries). Moreover, two libraries of the corresponding organs and blocks of the common parent B73 were assigned to each lane. Hence, the RNA samples of the three organs harvested from the same field block (≙ column) were assigned to three different lanes. In total, the 180 indexed libraries were loaded in nine lanes of three flow cells with 20 pooled libraries per lane. Cluster preparation and paired-end sequencing were performed according to the manufacturer's protocol (HiSeq 4000; Illumina).

From raw sequencing reads to read counts

We trimmed the 100-bp long paired-end sequencing reads by removing low-quality reads and adapter sequences using Trimmomatic version 0.36 (Bolger et al., 2014). In paired-end mode, we used the following settings: ILLUMINACLIP: 3:30:10, LEADING:3, TRAILING:3, MAXINFO:30:0.8, and MINLEN:40. We indexed B73v4 (Jiao et al., 2017; ftp://ftp.ensemblgenomes.org/pub/plants/release-42/fasta/zea_mays/dna/) using the alignment tool HiSat2 version 2.1.0 (Kim et al., 2015). We aligned the trimmed sequencing reads to the reference genome by providing a list of known splice sites extracted from the GTF annotation file of B73 (ftp://ftp.ensemblgenomes.org/pub/plants/release-42/gtf/zea_mays/). We set the minimum intron size to 20 and the maximum intron size to 60,000 and all other mapping parameters to default. In the following step, we marked and removed duplicates, that is, read pairs that have identical 5′-coordinates and orientations, from the mapped reads except the read pair having the highest sum of base qualities using the picard tool MarkDuplicates version 2.5.0 (http://broadinstitute.github.io/picard/command-line-overview.html). We obtained raw read counts for each protein-coding gene by using HTSeq-count version 0.10.0 using default settings (Anders et al., 2015). During this step, we discarded reads mapping at multiple positions of the genome and to overlapping gene models.

For subsequent comparative analyses between aboveground and belowground tissue expression patterns, we analyzed a sequencing dataset of three developmental stages of maize primary roots of the same genotype panel according to the same protocol described above (Baldauf et al., 2018); NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) accession numbers SRP110782).

Analysis of expression complementation

The gene activity status (on/off) was determined by thresholding normalized read counts. To adjust for artifactual read count differences across genes due to guanine-cytosine (GC) content and gene length, we fit a generalized additive model to log-transformed counts with GC content and log-transformed gene length as explanatory variables (Lithio and Nettleton, 2015). The R package *mgcv* (Wood, 2017) was used for parameter estimation. Given the estimated GC content and gene length effects, we obtained a predicted count for each gene. The inverse of this predicted count serves as a multiplicative gene-specific normalization factor. We also calculated sample-wise scale factors by using the trimmed mean of M-values (TMM) method to adjust for differences among library sizes (Robinson and Oshlack, 2010). Each raw count was then multiplied by the product of the appropriate gene-specific normalization factor and the TMM scale factor to obtain a normalized count. Given an organ type, we used the mean normalized count across all replicates of each genotype to represent the average expression level for each gene. For the root dataset, we used the 0.1 quantile of the nonzero average expression levels as the threshold for calling activity status. For the shoot dataset, the distribution of average expression levels varies across different organs; thus the 0.1 quantile of the nonzero average expression levels of each organ was used as an organ-specific threshold for calling activity status. We called a gene active if the average

expression level across all replicates is greater than the threshold and inactive otherwise for each genotype and organ type.

Regression analysis of the levels of MPH on the number of SPE genes

The baseline model has a fixed effect for blocks (4 levels) and genotypes (13 levels). As there are five plants measured per plot, the model has a random effect for plots. The residual error, therefore, assesses the within-plot variance among plants. The model was then extended to cover the regression of MPH on the number of SPE genes. Thus, we defined seven covariates for the seven parental lines. These covariates were initially all set to 0 for each observation. Then for observations on parental lines, the corresponding covariate for that line was set to 1. For observation on hybrids, the two covariates corresponding to the two parents were set to 0.5. Thus, collectively, these seven covariates model the effect of the per se performance of the parental lines and the mid-parent values of the hybrids. MPH was then modeled by a regression on the number of SPE genes. For this purpose, we set the number of SPE genes to zero for all parental lines. This was done to be able to include them in the overall model. Note, however, that the parental lines have no impact on the regression, because their effect is fully absorbed by the covariates for the parental line effects. As the MPH effects of the hybrids were not expected to fall on the regression line, we allowed for deviations from the regression by adding a random effect for the hybrids. This was implemented by fitting the random effect z*genotype, where z is a continuous dummy variable with $z = 0$ for parental lines and $z = 1$ for hybrid. This dummy variable acts as a switch that turns the random effect off for parental lines and turns it on for hybrids (Piepho et al., 2006). The analysis was implemented in the GLIMMIX procedure of SAS. The seven covariates for the parental lines were combined into an effect "parents" with six degrees of freedom using the "effect" statement of GLIMMIX. To plot the fitted model, we computed the product of the slope estimate multiplied by the number of SPE genes. Adding the predicted random effects for hybrids, we obtained the predicted MPH, which was then plotted against the number of SPE genes.

For the regression analysis of tissue-specific traits (i.e. size of the unpollinated ear, the mature ear, the flag leaf, and tassel), to the number of SPE genes, only the SPE genes were considered which were identified in the corresponding tissue dataset. In contrast, all SPE genes, which were identified in the three aboveground tissues, were used for the regression analysis of the remaining traits "plant height," "number of leaves," and "number of nodes." Similarly, all SPE genes identified in primary roots at three developmental stages (Baldauf et al., 2018) were used for the regression analysis of all aboveground phenotypic traits, to determine if SPEs manifested during very early plant development can be used to predict heterosis for aboveground traits later in development.

PCA for sample structure analysis

We performed a PCA on expression filtered (counts per million [CPM] >10 in $>70\%$ of all samples; Chen et al., 2016) normalized read counts of the active genes among all three aboveground organs using the function prcomp of the *stats* package in R (R version 4.0.1 (2020-06-06)). In addition, we included a previously published RNA-seq dataset covering three developmental stages of the maize primary root of the same genotype panel to analyze the aboveground organs in relation to the belowground dataset (Baldauf et al., 2018). In detail, we normalized the raw read counts of the active genes of both datasets by using the TMM method and transformed them into $\log_2$-CPM values (CPM; function calcNormFactors, Bioconductor package *edgeR_3.30.3*; Robinson and Oshlack, 2010). In addition, we filtered the active genes to keep only genes which have a CPM value >10 in $>70\%$ of all samples (function filterByExpr, Bioconductor package *edgeR_3.30.3*; Chen et al., 2016). Subsequently, we estimated the mean–variance trend within the filtered count data and assigned a precision weight to each observation (function voom, Bioconductor package *limma_3.44.3*; Law et al., 2014).

Gene co-expression network analysis and functional characterization

We performed WGCNA to identify clusters of highly connected genes and their correlation to phenotypic trait data (Zhang and Horvath, 2005). For each aboveground organ separately, we used only the genes displaying the highest variance (upper 75% percentile) among the expression filtered active genes from previous analyses for network construction using the *WGCNA* package in R (*WGCNA_version 1.69*; Langfelder and Horvath, 2008). We obtained co-expression modules for each organ dataset by using the automatic network construction function blockwiseModules with default settings, except the soft thresholding power (β) was set to 8 for ear, 6 for leaf, and 8 for the tassel dataset based on the scale-free topology fit index estimate. In addition, we set the network type to signed hybrid and TOMtype to signed. We selected a high module selection sensitivity by setting the corresponding parameter to 4 and merged modules with <30 genes into their closest larger neighboring module. We summarized each module by the first PC of the scaled module expression profiles, referred to as module eigengene. We set the parameter mergeCutHeight for merging the modules to 0.25, implying that modules whose eigengenes are correlated >0.75 will be merged. We visualized the identified network of eigengenes in heatmaps displaying the correlation to external phenotypic trait data and RNA-seq sample information.

To identify modules that are significantly associated with the measured phenotypic traits, we correlated module eigengenes with calculated plot means of the individual phenotypic traits using Pearson correlation in R. We calculated

student asymptotic p-values for each given correlation using the R function corPvalueStudent within the *WGCNA* package (*WGCNA_version 1.69*; Langfelder and Horvath, 2008).

To visualize the edges and nodes of the identified modules with significant module–trait correlations or enriched SPE gene expression patterns, we used the R package igraph and selected only the modules which display a correlation value $\geqslant |0.5|$ at a significance level <0.001 (***) for further analyses and graphical display. We removed edges with a weight <0.01 and unconnected nodes prior to plotting the network in R.

GO enrichment analysis and enrichment analysis of nonsyntenic genes

We used the Web-based agriGOv2 platform (Tian et al., 2017) to perform a GO enrichment analysis for the genes among the co-expression modules displaying significant sample and trait correlations. As reference for the singular enrichment analysis, we used the maize B73.AGPv4 annotation file of the maize-GAMER project (Wimalanathan et al., 2018). In addition to the singular enrichment analysis of each gene set, we performed a cross comparison among all enriched GO terms using the SEACOMPARE analysis tool of AgriGOv2 (Tian et al., 2017). Finally, for each gene set we summarized significantly enriched GO terms and reduced functional redundancies using the Web-based software REVIGO (Supek et al., 2011). We set the allowed similarity to 0.7. We determined the evolutionary origin of genes displaying SPE complementation or nonadditivity by comparing those genes with published lists of syntenic and nonsyntenic genes (Schnable, 2019). We applied Fisher's exact test at a significance level of $\alpha = 0.05$ to determine the distribution of nonsyntenic genes among those gene sets in comparison to all expressed genes that did not display the corresponding expression pattern.

Accession numbers

Raw sequencing data are deposited in the NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) under accession number PRJNA705557.

Supplemental data

The following materials are available in the online version of this article.

Supplemental Figure S1. Experimental designs of the field trial and loading of the flow cells for RNA-seq.

Supplemental Figure S2. Paired PC plots to visualize sample relations.

Supplemental Figure S3. Overlap of active genes between the three aboveground tissues among all 13 genotypes.

Supplemental Figure S4. Levels of MPH for diverse phenotypic traits and the six F_1-hybrids.

Supplemental Figure S5. Heatmaps displaying module–sample correlations in each aboveground tissue.

Supplemental Figure S6. Module–trait associations of all identified modules to phenotypic trait data.

Supplemental Table S1. Least-square means of each genotype and levels of MPH for all phenotypic traits.

Supplemental Table S2. Results of the regression analysis of the number of SPE genes on the level of MPH of diverse phenotypic traits.

Supplemental Table S3. Proportion of SPE_B and SPE_X genes among all genes assigned to each module identified in the unpollinated ear, flag leaf, and meiotic tassel.

Supplemental Table S4. GO enrichment analysis of modules enriched by SPE patterns.

Supplemental Data Set 1. Gene expression patterns identified among all protein-coding genes of B73v4 in all parent–hybrid triplets in the aboveground tissues unpollinated ear, flag, leaf, and meiotic tassel.

Acknowledgments

We thank Helmut Rehkopf and Christa Schulz (University of Bonn) for their support in propagating the genetic material for this study. We would like to thank the staff of the experimental field station Endenich of the University Bonn for crop cultivation.

Funding

This work was funded by the Deutsche Forschungsgemeinschaft (DFG) grants HO2249/9-3, HO 2249/18-1 and GRK2064 to F.H., DFG grant PI 377/19-2 to H.H.P. and the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) and the joint National Science Foundation (NSF)/NIGMS Mathematical Biology Program under award number R01GM109458 to D.N.

Conflict of interest statement. None declared.

References

Anders S, Pyl PT, Huber W (2015) HTSeq - a Python framework to work with high-throughput sequencing data. Bioinformatics **31**: 166–169

Austin DF, Lee M (1998) Detection of quantitative trait loci for grain yield and yield components in maize across generations in stress and nonstress environments. Crop Sci **38**: 1296–1308

Baldauf JA, Marcon C, Lithio A, Vedder L, Altrogge L, Piepho H-P, Schoof H, Nettleton D, Hochholdinger F (2018) Single-parent expression is a general mechanism driving extensive complementation of non-syntenic genes in maize hybrids. Curr Biol **28**: 431–437.e4

Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. J Stat Softw **67**: 1–51

Betrán FJ, Beck D, Bänziger M, Edmeades GO (2003) Secondary traits in parental inbreds and hybrids under stress and non-stress environments in tropical maize. F Crop Res **83**: 51–65

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**: 2114–2120

Botet R, Keurentjes JJB (2020) The role of transcriptional regulation in hybrid vigor. Front Plant Sci **11**: 410

Chen Y, Lun ATL, Smyth GK (2016) From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. F1000 Res **5**: 1438

Davidson RM, Hansey CN, Gowda M, Childs KL, Lin H, Vaillancourt B, Sekhon RS, de Leon N, Kaeppler SM, Jiang N, et al. (2011) Utility of RNA sequencing for analysis of maize reproductive transcriptomes. Plant Genome **4**: 191–203

Denancé N, Sánchez-Vallet A, Goffner D, Molina A (2013) Disease resistance or growth: the role of plant hormones in balancing immune responses and fitness costs. Front Plant Sci **4**: 155

Flint-Garcia SA, Buckler ES, Tiffin P, Ersoz E, Springer NM (2009) Heterosis is prevalent for multiple traits in diverse maize germplasm. PLoS One **4**: e7433

Gonzalez-Bayon R, Shen Y, Groszmann M, Zhu A, Wang A, Allu AD, Dennis ES, Peacock WJ, Greaves IK (2019) Senescence and defense pathways contribute to heterosis. Plant Physiol **180**: 240–252

Groszmann M, Gonzalez-Bayon R, Lyons RL, Greaves IK, Kazan K, Peacock WJ, Dennis ES (2015) Hormone-regulated defense and stress response networks contribute to heterosis in Arabidopsis F1 hybrids. Proc Natl Acad Sci USA **112**: E6397–406

Hansey CN, Vaillancourt B, Sekhon RS, de Leon N, Kaeppler SM, Buell CR (2012) Maize (*Zea mays* L.) genome diversity as revealed by RNA-sequencing. PLoS One **7**: e33071

Hirsch CN, Foerster JM, Johnson JM, Sekhon RS, Muttoni G, Vaillancourt B, Peñagaricano F, Lindquist E, Pedraza MA, Barry K, et al. (2014) Insights into the maize pan-genome and pan-transcriptome. Plant Cell **26**: 121–135

Hochholdinger F, Baldauf JA (2018) Heterosis in plants. Curr Biol **28**: R1089–R1092

Hochholdinger F, Hoecker N (2007) Towards the molecular basis of heterosis. Trends Plant Sci **12**: 427–32

Hochholdinger F, Woll K, Sauer M, Dembinsky D (2004) Genetic dissection of root formation in maize (*Zea mays*) reveals root-type specific developmental programmes. Ann Bot **93**: 359–368

Hoecker N, Keller B, Piepho HP, Hochholdinger F (2006) Manifestation of heterosis during early maize (*Zea mays* L.) root development. Theor Appl Genet **112**: 421–429

Huot B, Yao J, Montgomery BL, He SY (2014) Growth-defense tradeoffs in plants: a balancing act to optimize fitness. Mol Plant **7**: 1267–1287

Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, Campbell MS, Stein JC, Wei X, Chin C, et al. (2017) Improved maize reference genome with single-molecule technologies. Nature **546**: 524–527

Jin M, Liu H, He C, Fu J, Xiao Y, Wang Y, Xie W, Wang G, Yan J (2016) Maize pan-transcriptome provides novel insights into genome complexity and quantitative trait variation. Sci Rep **6**: 18936

Jones DF (1917) Dominance of linked factors as a means of accounting for heterosis. Genetics **2**: 466–479

Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. Nat Methods **12**: 357–360

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics **9**: 559

Larièpe A, Mangin B, Jasson S, Combes V, Dumas F, Jamin P, Lariagon C, Jolivot D, Madur D, Fiévet J, et al. (2012) The genetic basis of heterosis: multiparental quantitative trait loci mapping reveals contrasted levels of apparent overdominance among traits of agronomical interest in maize (*Zea mays* L.). Genetics **190**: 795–811

Law CW, Chen Y, Shi W, Smyth GK (2014) voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol **15**: R29

Li Z, Coffey L, Garfin J, Miller ND, White MR, Spalding EP, De Leon N, Kaeppler SM, Schnable PS, Springer NM, et al. (2018) Genotype-by-environment interactions affecting heterosis in maize. PLoS One **13**: e0191321

Li Z, Zhou P, Della Coletta R, Zhang TF, Brohammer AB, O'Connor CH, Vaillancourt B, Lipzen A, Daum C, et al. (2021) Single-parent expression drives dynamic gene expression complementation in maize hybrids. Plant J **105**: 93–107

Lithio A, Nettleton D (2015) Hierarchical modeling and differential expression analysis for RNA-seq experiments with inbred and hybrid genotypes. J Agric Biol Environ Stat **20**: 598–613

Lorenz A, Hoegemeyer T (2013) The phylogenetic relationships of US maize germplasm. Nat Genet **45**: 844–845

Marcon C, Paschold A, Malik WA, Lithio A, Baldauf JA, Altrogge L, Opitz N, Lanz C, Schoof H, Nettleton D, et al. (2017) Stability of single parent gene expression complementation in maize hybrids upon water deficit stress. Plant Physiol **173**: 1247–1257

Marroni F, Pinosio S, Morgante M (2014) Structural variation and genome complexity: is dispensable really dispensable? Curr Opin Plant Biol **18**: 31–36

Mehta H, Sarkar KR (1992) Heterosis for leaf photosynthesis, grain yield and yield components in maize. Euphytica **61**: 161–168

Melchinger AE, Gumber RK (1998) Overview of heterosis and heterotic groups in agronomic crops. *In* KR Larnkey, JE Staub, eds, Concepts Breed Heterosis Crop Plants. CSSA Spec. Publ. 25, Madison, WI, pp 29–44

Miller M, Song Q, Shi X, Juenger TE, Chen ZJ (2015) Natural variation in timing of stress-responsive gene expression predicts heterosis in intraspecific hybrids of Arabidopsis. Nat Commun **6**: 7453

Paschold A, Jia Y, Marcon C, Lund S, Larson NB, Yeh C-T, Ossowski S, Lanz C, Nettleton D, Schnable PS, et al. (2012) Complementation contributes to transcriptome complexity in maize (*Zea mays* L.) hybrids relative to their inbred parents. Genome Res **22**: 2445–2454

Paschold A, Larson NB, Marcon C, Schnable JC, Yeh CT, Lanz C, Nettleton D, Piepho HP, Schnable PS, Hochholdinger F (2014) Nonsyntenic genes drive highly dynamic complementation of gene expression in maize hybrids. Plant Cell **26**: 3939–3948

Paschold A, Marcon C, Hoecker N, Hochholdinger F (2010) Molecular dissection of heterosis manifestation during early maize root development. Theor Appl Genet **120**: 383–388

Piepho HP, Williams ER, Fleck M (2006) A note on the analysis of designed experiments with complex treatment structure. HortScience **41**: 446–452

Piepho HP (2005) Optimal allocation in designs for assessing heterosis from cDNA gene expression data. Genetics **171**: 359–364

Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol **11**: R25

Schnable JC (2019) Grass syntenic gene list sorghum v3 maize v3/4 with teff and oropetium v2. https://figshare.com/articles/dataset/Grass_Syntenic_Gene_List_sorghum_v3_maize_v3_4_with_teff_and_oropetium_v2/7926674/1

Schnable JC (2015) Genome evolution in maize: from genomes back to genes. Annu Rev Plant Biol **66**: 329–343

Schnable JC, Lyons E (2011) Comparative genomics with maize and other grasses: from genes to genomes! Maydica **56**: 183–200

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol **7**: 3

Shull GH (1948) What is "heterosis"? Genetics **33**: 439–446

Springer NM, Ying K, Fu Y, Ji T, Yeh CT, Jia Y, Wu W, Richmond T, Kitzman J, Rosenbaum H, et al. (2009) Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. PLoS Genet **5**: e1000734

Stelpflug SC, Rajandeep S, Vaillancourt B, Hirsch CN, Buell CR, Leon N De, Kaeppler SM (2016) An expanded maize gene expression atlas based on RNA-sequencing and its use to explore root development. Plant Genome **9**: 314–362

Supek F, Bošnjak M, Škunca N, Šmuc T (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One **6**: e21800

Swanson-Wagner RA, Eichten SR, Kumari S, Tiffin P, Stein JC, Ware D, Springer NM (2010) Pervasive gene content variation

and copy number variation in maize and its undomesticated progenitor. Genome Res **20**: 1689–1699

Tian T, Liu Y, Yan H, You Q, Yi X, Du Z, Xu W, Su Z (2017) agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res **1**: 1–8

Troyer FA (2001) Temperate corn - Background, behaviour, and breeding. *In* R A. Hallauer, ed, Specialty Corns. CRC Press, Boca Raton, FL, pp 393–466

Wang X, Cao H, Zhang D, Li B, He Y, Li J, Wang S (2007) Relationship between differential gene expression and heterosis during ear development in Maize (*Zea mays* L.). J Genet Genomics **34**: 160–170

Wimalanathan K, Friedberg I, Andorf CM, Lawrence-Dill CJ (2018) Maize GO annotation - Methods, evaluation, and review (maize-GAMER). Plant Direct **2**: 1–15

Wood SN (2017) Generalized Additive Models: An Introduction with R, 2nd ed. Chapman and Hall/CRC, London

Yang J, Mezmouk S, Baumgarten A, Buckler ES, Guill KE, McMullen MD, Mumm RH, Ross-Ibarra J (2017) Incomplete dominance of deleterious alleles contributes substantially to trait variation and heterosis in maize. PLoS Genet **13**: e1007019

Yi Q, Liu Y, Hou X, Zhang X, Li H, Zhang J, Liu H, Hu Y, Yu G, Li Y, et al. (2019) Genetic dissection of yield-related traits and mid-parent heterosis for those traits in maize (*Zea mays* L.). BMC Plant Biol **19**: 392

Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol **4**: 17

Zhang X, Ma C, Wang X, Wu M, Shao J, Huang L, Yuan L, Fu Z, Li W, Zhang X, et al. (2021) Global transcriptional profiling between inbred parents and hybrids provides comprehensive insights into ear-length heterosis of maize (*Zea mays*). BMC Plant Biol **21**: 118