Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

Review

Allelic variation and heterosis in maize: How do two halves make more than a whole?

Nathan M. Springer[1] and Robert M. Stupar

Cargill Center for Microbial and Plant Genomics, Department of Plant Biology, University of Minnesota, Saint Paul, Minnesota 55108, USA

In this review, we discuss the recent research on allelic variation in maize and possible implications of this work toward our understanding of heterosis. Heterosis, or hybrid vigor, is the increased performance of a hybrid relative to the parents, and is a result of the variation that is present within a species. Intraspecific comparisons of sequence and expression levels in maize have documented a surprisingly high level of allelic variation, which includes variation for the content of genic fragments, variation in repetitive elements surrounding genes, and variation in gene expression levels. There is evidence that transposons and repetitive DNA play a major role in the generation of this allelic diversity. The combination of allelic variants provides a more comprehensive suite of alleles in the hybrid that may be involved in novel allelic interactions. A major unresolved question is how the combined allelic variation and interactions in a hybrid give rise to heterotic phenotypes. An understanding of allelic variation present in maize provides an opportunity to speculate on mechanisms that might lead to heterosis. Variation for the presence of genes, the presence of novel beneficial alleles, and modified levels of gene expression in hybrids may all contribute to the heterotic phenotypes.

It is often argued that a basic understanding of a process is critical to its use and manipulation. Arguably, one of the most notable exceptions to this idea is the phenomenon of heterosis. Heterosis, or hybrid vigor, has been the subject of intense research and speculation for well over a century; however, the basic mechanisms that cause or contribute to heterosis remain unclear (Coors and Pandey 1999). Despite this lack of understanding, breeders have quite successfully manipulated heterosis to increase the vigor of many domesticated species. In maize, it is estimated that the use of hybrids and heterosis increases yields ~15% per annum (Duvick 1999).

Heterosis refers to the phenomenon in which the hybrid F_1 offspring exhibit phenotypic characteristics that are superior to the mean of the two parents (mid-parent heterosis), or the better of the two parents (better parent heterosis) (Fig. 1). While mid-parent heterosis is scientifically interesting, it has relatively little economic importance. Better parent heterosis is the underlying rationale for the widespread use of hybrids in many agricultural species, and we focus our discussion on this phenomenon. The level of heterosis can be quantified for specific traits. "Heterosis per se" for a specific trait is quantified as the phenotype of the hybrid minus the phenotype of the better parent; subsequently, the "percentage of heterosis" is calculated as the heterosis per se divided by the better of the parental phenotypes. Often, the parents of heterotic offspring are inbred. In this case, the quantification of heterosis reflects both hybrid vigor and a recovery from inbreeding depression.

Heterosis has been used in the breeding and production of many crop and animal species (Janick 1998; Melchinger and Gumber 1998). An early application of heterosis was breeding mules, which are derived from crossing a female horse and a male donkey (Goldman 1998). As pointed out by Troyer (2006), there are useful similarities in the examples of mules and hybrid corn. Both mules and hybrid corn exhibit superior phenotypes and stress tolerance relative to their parents. Additionally, farmers showed a willingness to purchase hybrid corn and mules despite obvious drawbacks: these products are added expenses and neither produce useful offspring, thus requiring the farmer to purchase a new organism each generation.

It is important to note that the application and use of hybrids in different species are influenced by several factors. For instance, if the degree of heterosis is relatively low in a given species, then it may not be cost-effective to use hybrids for commercial production. The efficient development of hybrids may be limited by the species mating system, as self-pollination and/or controlled crosses are difficult or impossible to conduct in many species. An efficient F_1 hybrid production method is essential for hybrid commercialization. In fact, the early maize inbred varieties had extremely low seed yield, leading to the use of double cross hybrids until inbred varieties with higher seed yield could be developed (Duvick 2001).

Heterosis is evident not only in artificially selected populations, but also may be observed in natural populations (Mitton 1998; Hansson and Westerberg 2002). Allelic frequencies in a random sample of coniferous trees were in agreement with Hardy-Weinberg expectations; however, an excess of heterozygotes was observed when only the mature, oldest, or largest trees were sampled (Mitton and Jeffers 1989). Furthermore, a study of Pinyon pines suggested that heterozygotes are more resistant to herbivory pressure (Mopper et al. 1991). The exact mechanism that leads to the enhanced performance of the heterozygotes in these tree species has not been determined, but it is possible that heterosis is an important factor in fitness for many organisms.

Maize provides an excellent system for the study and application of heterosis. A wide range of natural genetic diversity has been captured in the current maize germplasm (Flint-Garcia et al. 2005; Wright et al. 2005; Troyer 2006). Maize is relatively easy to self- or cross-pollinate, which has enabled the development of both diverse inbreds and many hybrids for evaluation. Much of

[1]**Corresponding author.**
E-mail springer@umn.edu; fax (612) 625-1738.
Article published online before print. Article and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.5347007.

17:264–275 ©2007 by Cold Spring Harbor Laboratory Press; ISSN 1088-9051/07; www.genome.org

PIONEER 2017
Inari Agric. Inc. v. Pioneer Hi-Bred Int'l, Inc.
PGR2024-00025
1/13

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

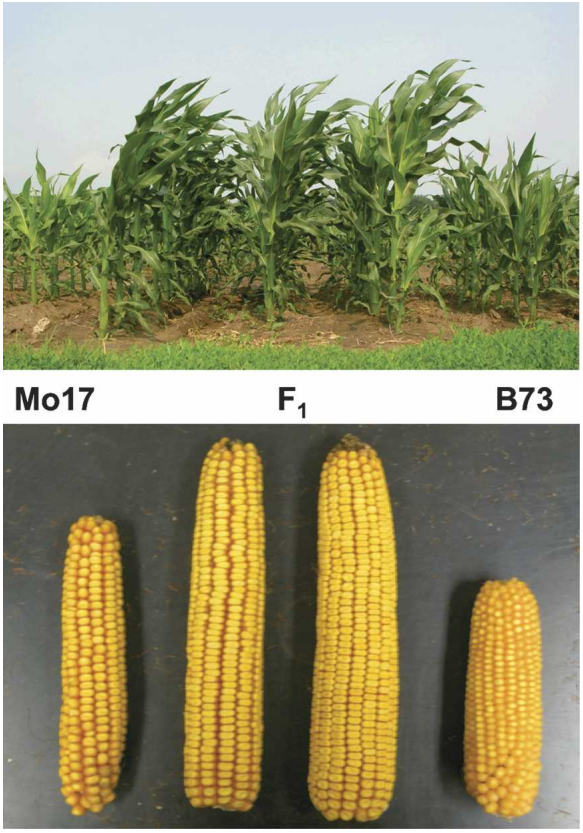

Figure 1. Phenotypic heterosis in the B73 × Mo17 hybrid. Representative B73, Mo17, and F_1 hybrid ears and plants are shown. Note the increased size of the two hybrid ears and three hybrid rows relative to the two inbred parents (*left*, Mo17; *right*, B73).

this review focuses on the heterosis derived from crossing maize inbred line B73 with inbred line Mo17 (Fig. 1). The B73 × Mo17 hybrid was widely grown commercially in the United States during the 1970s, with >15 million bags of seed sold, sufficient to plant ~45 million acres (Troyer 2006). While there are many maize hybrids that exhibit high levels of heterosis, this particular hybrid provides an excellent example of heterosis involving two lines that are relatively well-characterized in terms of genetic mapping, gene expression, and genome organization.

The relative measurements for several traits of the B73 and Mo17 parental inbred lines and the hybrid are listed in Table 1 (Zanoni and Dudley 1989; Auger et al. 2005b). Note that the quantitative measurements of heterosis vary significantly for different traits. The variation in levels of heterosis among traits leads to difficulties in accurately quantifying the amount of overall heterosis. There is also variation in the relative level of heterosis for different traits between different hybrids. For example, one maize hybrid may exhibit significant heterosis for plant height but have low heterosis for grain yield, while a second hybrid may exhibit high heterosis for grain yield but very low heterosis for height (Zanoni and Dudley 1989; data not shown). This variation suggests that the same set of genes does not control all heterotic responses. Additionally, heterosis does not simply result from the overall genetic diversity within a hybrid, but is likely a reflection of diversity at specific, important genes that contribute to a particular trait. This view is supported by the ability to map QTLs that contribute to heterosis for individual traits in maize and rice (Stuber et al. 1992; Xiao et al. 1995; Li et al. 2001; Luo et al. 2001; Lu et al. 2003).

In the simplest view, the reverse of heterosis is inbreeding depression, in which progressive self-pollination or sibling matings reduce the genome-wide heterozygosity and overall fitness of an organism. Inbreeding depression is likely caused by the fixation of deleterious alleles within a lineage. Plant breeders have created an "unnatural" inbred state in maize, a naturally outcrossing species, which may help explain why maize exhibits relatively high heterosis when compared to other species. The inbreeding depression in inbred maize lines is alleviated when a hybrid is formed, and the plants exhibit superior characteristics. A fundamental question is whether the reversal of the effects of inbreeding in the F_1 hybrid (via complementation of deleterious alleles with "good" alleles from the other parent) is sufficient to explain heterosis, or whether other mechanisms are at work. Two hypotheses, dominance and overdominance, have been proposed to explain heterosis (Text Box 1).

In this review, we focus on three key questions that address heterosis from a genomics perspective: (1) What allelic differences exist between heterotic parental lines in terms of types of variation and prevalence of variation? (2) How do these allelic variants interact in the F_1 hybrid organism? (3) How does the interaction of allelic variants result in heterotic phenotypes?

The term "allelic variation" (broadly defined here as the sequence or regulatory differences found in different parental genotypes) is used in each of the three questions. While we discuss each of these questions in the context of maize, it is anticipated that similar mechanisms apply to other organisms.

Allelic variation in maize

The current era of genomic sequencing and global gene expression analysis has provided a wealth of information about intraspecific allelic variation. Intraspecific allelic variation can include sequence changes, structural changes in the genome, altered expression levels, and epigenetic changes. At some level, heterosis is the result of variation between the parental lines, since in the

Table 1. Phenotypes and heterosis in B73 relative to Mo17

Trait	B73	Mo17	Hybrid	Heterosis per se[a]	% heterosis[a]
Yield (mg/ha)[b]	6.32	4.06	10.41	4.1	64.7%
Seed number[c]	430	271	628	198.0	46.0%
Seed weight (g/ear)[d]	115	98	232	117.0	101.7%
Height (cm)[d]	173.3	169	211.1	37.8	21.8%
Total nodes[d]	14.6	13.3	15.1	0.5	3.4%
Ear node[d]	8.9	8.4	9.2	0.3	3.4%
Leaf length (cm)[d]	75.3	62.7	87	11.7	15.5%
Leaf width (mm)[d]	90.1	84.2	106.4	16.3	18.1%
Days to first silk[d]	70.1	77.1	68.05	2.1	2.9%
Days to first anther[d]	68.9	71.1	65.95	3.0	4.3%
Tassel branches[d]	9.3	6.9	9.75	0.4	4.8%
Ear length (cm)[d]	12.7	14.3	21.2	6.9	48.3%
11-d seedling height (cm)[c]	32.4	27.3	41	8.6	26.5%
11-d seedling biomass (g)[c]	0.42	0.32	0.71	0.3	69.0%

[a]Better-parent heterosis.
[b]Zanoni and Dudley (1989).
[c]Unpublished data, N.M. Springer and R.S. Stupar.
[d]Auger et al. (2005a).

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

Text Box 1. Dominance and overdominance models of heterosis

The genetic basis of heterosis has been debated for nearly a century without a clear resolution. The two main hypotheses advanced to explain this phenomenon are dominance (Davenport 1908; Jones 1917) and overdominance (East 1908; Shull 1908). These models have survived with various modifications and interpretations as the methods and specialties of biology have changed (Crow 1999; Birchler et al. 2003). The dominance model of heterosis posits that each of the inbred lines contains slightly deleterious alleles that reduce their fitness. The hybrid will benefit from the complementation of these deleterious alleles and will display a superior phenotype. The overdominance model suggests that the heterozygous combination of alleles at a given locus is phenotypically superior to either of the homozygous combinations for that locus; contributions of overdominant loci thereby result in a superior hybrid.

Genetic linkage causes major difficulties in resolving dominant from overdominant action on phenotypes. A locus may display apparent overdominant action that is actually the result of pseudo-overdominance, which occurs when two repulsion-phase-linked loci both contribute via dominant mechanisms (Stuber et al. 1992; Graham et al. 1997). Bingham (1998) provided a discussion of the role of chromosomal blocks and the ability to resolve overdominance from pseudo-overdominance. In maize, it is likely that only 30 chromosomal blocks (a segment of a chromosome derived from one parent) are transmitted (~2 or 3 crossovers per chromosome pair). Given this limited resolution of recombination at the individual level, it is anticipated that many repulsion-phase linkages will not be broken, resulting in pseudo-overdominance. The use of advanced populations in maize derived from random intermating provided strong evidence to support widespread pseudo-overdominance, as the evidence for overdominant loci was reduced over successive generations (for reviews, see Bingham 1998; Crow 1999).

Although the current view favors dominant allelic action as the major contributor to heterosis (Coors and Pandey 1999; Crow 1999), there are several lines of evidence that suggest that mechanisms beyond simple complementation may be important in heterosis. The absence of a decline in the magnitude of heterosis (Duvick 2001), the progressive heterosis in tetraploids, and the rapid rate of inbreeding depression in tetraploids have been cited as factors that suggest that dominance may be insufficient to completely explain heterosis (Birchler et al. 2003). In addition, the existence of a locus with overdominant phenotypic action has been demonstrated in maize (Hollick and Chandler 1998). A recent study using introgression lines documented the contribution of dominance and overdominance to heterosis in tomato (Semel et al. 2006). This study found significant contributions of overdominant action to reproductive traits but not other traits. However, reproductive traits were also influenced by a larger number of QTL than other traits, which suggests a potential contribution of pseudo-overdominance.

The dominance/overdominance debate becomes even more nuanced when contributions of epistasis are considered. Epistasis is classically defined as interactions between genes at two (or more) loci affecting the phenotypic expression of a trait. With regard to heterosis, a dominant-acting locus may interact with an overdominant-acting locus to further heterotic gain. Therefore, the genetic background and allelic interactions therein can have an effect on the heterotic contributions of individual loci.

absence of variation (inbreds), there is no heterosis. Our assumptions of the nature and frequency of allelic variation can limit how we consider heterosis (Crow 1999; Phillips 1999).

Allelic diversity in maize genic sequences

One of the most common approaches toward documenting allelic diversity is to compare the sequence of genic regions (including coding regions, introns, untranslated regions, and single-copy DNA surrounding genes) from multiple strains or varieties in order to identify variation. This variation can then be used for mapping or association studies. An investigation of randomly selected sequences in the maize inbred B73 relative to inbred Mo17 found that, on average, indel polymorphisms occur every 309 bp and SNPs occur every 79 bp (Vroh Bi et al. 2005). The analysis of 300–500 bp amplicons found that 44% of the sequences contained at least one polymorphism in B73 relative to Mo17. In general, it is estimated that there is one polymorphism every 100 bp in any two randomly chosen maize inbred lines (Tenaillon et al. 2001; Ching et al. 2002). Sequence diversity data have been evaluated for several hundred diverse inbred lines of maize at >3000 genic sites (Zhao et al. 2006). These data provide extensive information about the SNP and indel frequencies of maize alleles. Collectively, these studies indicate that maize has a relatively high level of sequence polymorphism compared to many other species. For example, the level of sequence diversity in genic sequences within maize is estimated to be higher than the level of diversity between humans and chimpanzees (Buckler et al. 2006).

Structural diversity in maize

Structural genome diversity involves alterations in DNA sequence beyond SNPs or small indels. This would include large-scale chromosomal differences, altered location of genes or repetitive elements, or differences in the presence of sequences in one inbred relative to another. Large-scale genome differences between different maize inbred lines were first identified through cytogenetics studies. Barbara McClintock and others analyzed heterochromatic knob (highly condensed, tandem repeat regions) content and size to characterize genome variation in maize (Brown 1949; McClintock et al. 1981; Adawy et al. 2004). Recent studies have documented differences in the content for several classes of repetitive DNA between maize inbreds at the chromosomal level (Kato et al. 2004). Flow cytometry studies have also documented significant variation in the size of the maize genome between different inbred lines (Laurie and Bennett 1985; Lee et al. 2002).

Sequence-based methodologies have documented maize genome structural diversity at a higher resolution (for reviews, see Buckler et al. 2006; Messing and Dooner 2006). BAC-based studies focusing on the diversity of allelic regions among different maize inbred lines identified a surprising level of intraspecific structural diversity at several loci (Fig. 2). Fu and Dooner (2002) sequenced BAC contigs that contain the *bz* gene from the inbred B73 and McC genotypes and, as expected, documented numerous SNP and short indel polymorphisms within the genes present in this region. In addition, they found a surprising level of structural diversity in the two regions that resulted in extensive sequence nonhomologies. This structural diversity includes variation in repetitive elements and variation in the presence of genic fragments.

Seven conserved genes were found in the *bz* region of chromosome 9 for both B73 and McC; however, an additional four gene-like fragments were found only in the McC haplotype (Fu and Dooner 2002). Initially these fragments were thought to be genes specific to the McC genotype. However, further investigation found that these fragments are actually partial gene segments that were likely the result of two transposition events mediated by *Helitron* transposable elements (Lai et al. 2005). The

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

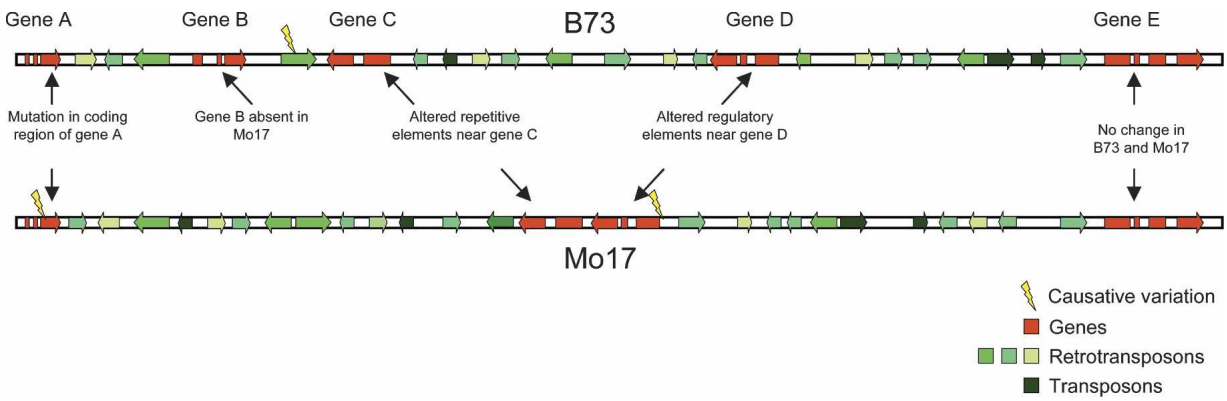

Figure 2. Potential examples of allelic variation in maize. A hypothetical small chromosomal region is diagrammed for B73 and Mo17. The genes (red) are separated by clusters of transposons (black) and several different classes of retrotransposons (various shades of green). There are likely to be numerous SNPs and small indels within this region. The composition of repetitive sequences also shows significant variation between the two inbreds. The specific location of a causative change that has resulted in functional allelic variation is indicated by the yellow lightening bolts. Gene A is an example of a variant in which the proteins are different. Gene B is a nonshared sequence that is only present in B73 at this locus. Gene C is more highly expressed in Mo17 than in B73 because of differences in the repetitive elements surrounding this gene in B73 and Mo17. Gene D is expressed only in B73 and not in Mo17 because of altered sequence in nearby regulatory regions. Gene E shows no functional variation in B73 relative to Mo17.

partial fragments present at the McC *bz* locus on chromosome 9 are derived from full-length gene sequences that are present on chromosome 5 in both B73 and McC (Lai et al. 2005). *Helitron* transposons, which use a rolling-circle method of replication, exhibit limited primary sequence signatures and can be difficult to identify using computational approaches (Kapitonov and Jurka 2001). The exact mechanism through which these elements can "capture" and transpose genic fragments is not well-understood (Brunner et al. 2005b; Lai et al. 2005).

The B73 and McC genotypes also demonstrated structural variation for repetitive elements present in this region (Fu and Dooner 2002). Only one short fragment from a repetitive *Zeon1* element is conserved in both genotypes. Otherwise, the transposable elements present in McC and B73 are completely distinct in terms of their relative positions. The McC haplotype contains six full-length LTR retrotransposons and six partial elements that are not conserved in the B73 locus, while the B73 haplotype contains five full-length LTR retrotransposons and four partial elements that are not present in the McC locus (Fu and Dooner 2002). In many cases, the B73 and McC alleles of the same gene were surrounded by distinct repetitive elements that may contribute to differences in the chromatin neighborhood for the two alleles. A recent follow-up study compared the genomic sequence structure of the *bz* locus in eight different maize lines representing a wide range of geographic and genetic backgrounds (Wang and Dooner 2006). A tremendous degree of variation in genomic sequence structure was identified among these lines; the authors concluded that, on average, any two of the eight *bz* haplotypes share just 50% of their sequences (Wang and Dooner 2006).

Several other maize loci that were sequenced in multiple inbred lines also exhibit structural diversity for genic fragments and repetitive elements. The tandemly repeated *zein1C* gene family displays variation between maize inbred lines in terms of copy number and expression (Song and Messing 2003). The sequencing of four genomic loci in B73 and Mo17 identified 45 genes that were shared between genotypes, while another 27 genic fragments (which are likely pseudogenes) were only present in either one of the two inbred lines (Brunner et al. 2005a). These four loci also included a set of 27 LTR retroelements that are present at collinear locations in both B73 and Mo17, while another 62 LTR retroelements were only present in either one of the two inbred lines.

The genome-wide prevalence of nonshared genic fragments was investigated using BAC library hybridization methods (Morgante et al. 2005). Hybridization of short genic probes to fingerprinted BAC libraries from B73 and Mo17 resulted in mapping ~20,500 genes, of which 80% were present in both inbreds, 11% were specific to B73, and 9% were specific to Mo17. Applying these frequencies to the anticipated genome size of maize suggests that the B73 and Mo17 genomes may contain ~10,000 nonshared genic fragments; a subset of which may be functional.

There is evidence that mechanisms other than *Helitron*-mediated transposition can lead to the presence of nonshared genic fragments in different varieties of a species (Bennetzen 2005). For instance, other classes of transposons can modify genic content. *Mutator*-like transposons can "capture" genic sequences and move these sequences within the maize genomes (Bureau et al. 1994; Jin and Bennetzen 1994; Bennetzen 2005). The abundance of the *Mutator*-like elements that carry gene fragments, termed Pack-MULEs, is not known in maize; however, the completion of the rice genome has allowed for a comprehensive analysis of these elements in rice (Jiang et al. 2004). Fragments of >1000 genes have been captured and sometimes duplicated by Pack-MULEs, resulting in the presence of >3000 Pack-MULEs in rice. There is EST evidence for the expression of ~5% of the Pack-MULEs, and some of these expressed sequences may encode functional peptides, although most contain numerous stop codons. There is also evidence for the presence of large numbers of tandem duplications in the maize genome, including some examples of tandem duplications that are present only in some inbred lines (Emrich et al. 2007). Duplicated genes and transposed gene fragments, like those discussed above, are likely to show variation within a species such that specific duplications or Pack-MULEs will be present within one inbred and absent in others.

The presence of nonshared genic sequences or retrotransposons might effect phenotypic variation through a variety of mechanisms. The majority of examples of the nonshared genic sequences that have been characterized to date are partial gene

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

fragments. It remains possible that some of these nonshared genic sequences will include full-length genes or novel genes created by shuffling of exons from multiple genes (Brunner et al. 2005b). In fact, there is evidence for expression of some of these nonshared genes (Brunner et al. 2005b; Morgante et al. 2005). The expression level and pattern of these fragments are likely to be distinct from those of the original gene because the newly transposed gene copy often does not include the full promoter or regulatory sequences. In some cases, the transcription and translation of a transposed partial ORF may result in an aberrant protein that has a dominant-negative effect that inhibits the function of the original full-length protein.

The variation in maize repetitive elements may influence the expression level of nearby genes, and in some cases, this influence can extend over large regions. For example, most alleles of the maize transcription factor *B* have relatively small, closely linked, regulatory sequences. However, the transcription level of the *B-I* allele can be influenced in *cis* by a set of tandem repeats that are located >100 kb upstream of the coding region (Stam et al. 2002). Additionally, the *tb1* locus, which influences maize architecture, is influenced in *cis* by a region of repetitive elements >50 kb upstream of the gene (Clark et al. 2006). These examples suggest that in some cases, repetitive sequences may affect the expression level of maize genes, ultimately resulting in altered phenotypes. However, it is unknown whether these examples represent rare cases of long-distance effects of repetitive sequences, or if this phenomenon is common in maize gene regulation.

Expression diversity in maize

Intraspecific allelic variation also includes variation in gene expression. Several recent studies have documented expression diversity in maize. Analysis of gene expression levels using an Agilent long oligonucleotide microarray platform found evidence for the differential expression of ~5% of genes in the maize inbred lines A619 and W23 at several developmental stages (Ma et al. 2006). A cDNA microarray platform used by Swanson-Wagner et al. (2006) identified ~10% of genes as being differentially expressed in B73 and Mo17 seedlings. Affymetrix microarray analysis also found prevalent differential expression between B73 and Mo17 in seedling, immature ear, and embryo tissues (Stupar and Springer 2006).

The Affymetrix microarray study suggested that ~2.5% of the genes present on the array were detected in only one of the two transcriptomes, either B73 or Mo17 (Stupar and Springer 2006). Validation of these results provided evidence that many of these differences in B73–Mo17 transcriptome content are not the result of differences in genomic content. Instead, these presence–absence transcriptome differences appear to be the result of differentially expressed genes that are present in the genomes of both B73 and Mo17.

The expression diversity observed in maize could be the result of *cis*-acting variation at each of the differentially expressed genes or the result of variation at a small number of *trans*-acting loci that have downstream regulatory effects. *Cis*-acting variation can be the result of alterations in regulatory sequences (i.e., enhancers and promoters), sequence changes that affect the RNA stability, or heritable variation in chromatin structure. *Trans*-acting variation can be the result of quantitative or qualitative variation in a factor that influences the expression of the gene, such as a transcription factor. The prevalence of *cis*- and *trans*-regulatory variation can be assessed using expression quantitative trait loci (eQTL) (Cheung and Spielman, 2002) and allele-specific expression (ASE) (Wittkopp et al. 2004) analyses. Numerous studies in animal systems using both eQTL (Monks et al. 2004; Morley et al. 2004; Cheung et al. 2005; Doss et al. 2005; Pastinen et al. 2005; Stranger et al. 2005) and ASE (Cowles et al. 2002; Yan et al. 2002; Bray et al. 2003; Lo et al. 2003; Pastinen et al. 2004; Wittkopp et al. 2004) have found that a significant amount of expression diversity is the result of *cis*-acting variation.

There is evidence for prevalent *cis*-acting regulatory variation in maize. An eQTL study found that 80% of the eQTL with a LOD score >7.0 mapped to the same physical location as the differentially expressed gene, indicating that allelic *cis*-variation may be causing regulatory differences for many maize genes (Schadt et al. 2003). Guo et al. (2004) used ASE to study the relative expression of two alleles in maize F_1 hybrids. Both alleles present in an F_1 hybrid have access to an identical set of *trans*-acting factors, thus unequal, or biased, allelic expression suggests *cis*-acting variation between alleles. Frequent allelic expression bias in F_1 hybrids was observed (11/15 genes), suggesting that *cis*-variation is present in many maize genes (Guo et al. 2004). These findings were supported by another ASE study that found allelic *cis*-variation contributing to inbred expression differences in 46 of 53 maize genes (Stupar and Springer 2006).

In addition to structural and expression diversity, there are additional sources of diversity within maize. Variation in sense–antisense transcription, allelic variation in DNA methylation patterns, and allelic variation in chromatin structure are all topics that have not been extensively addressed to date. One study of the targets of CpNpG methylation in maize found substantial variation for the epialleles present in B73 and Mo17 (I. Makarevitch and N.M. Springer, unpubl.). In addition, studies of polyploids have found evidence for widespread epigenetic alterations that can lead to phenotypic variation (Osborn et al. 2003). The epigenetic diversity within a species may be particularly important in contributing to overdominance. For example, one of the few examples documenting overdominant gene action for a single locus involves an epiallele at the *Pl* locus in maize (Hollick and Chandler, 1998). Phillips (1999) presented an excellent discussion of the evidence for de novo variation in inbred lines and how this variation may contribute to heterosis.

Interactions of alleles in hybrids

The allelic variation described above is brought together in the hybrid organism. The allelic combinations present in a hybrid may result in interactions that alter expression profiles, new protein–protein interactions, or epistatic interactions. The studies of gene expression in inbreds and hybrids can be divided according to whether they have found substantial evidence for frequent nonadditive expression patterns compared to the expression levels in the inbred parents (see Text Box 2 for a primer on the terminology and concepts associated with gene expression in inbreds relative to hybrids). The concept of transcriptional epistasis and the evidence for this phenomenon in hybrid maize are also discussed.

Evidence for allelic interactions that lead to novel hybrid expression patterns

Several studies have found evidence for novel expression patterns in hybrids relative to inbred parents such that the hybrid

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

Text Box 2. Describing the gene expression levels in inbreds and hybrids

Hybrid expression patterns can be classified as additive or nonadditive relative to the expression patterns observed in the inbred parents. An additive expression pattern occurs when the hybrid expression level is equal to the average of the two parents; for this reason, the term additive is often used interchangeably with the term mid-parent. Nonadditive expression patterns occur whenever the hybrid expression level deviates from the mid-parent value.

The terms "dominant" and "overdominant" are often used to describe hybrid gene expression levels compared to parental levels. The usage of these terms can lead to confusion with the quantitative genetics models of dominance or overdominance (see Text Box 1). There is not a strict relationship between the two concepts, as overdominant expression could lead to dominant phenotypes, or dominant expression might lead to overdominant phenotypes. Therefore, we have purposefully not used these terms to describe inbred-hybrid expression patterns. Instead, a systematic terminology to describe hybrid gene expression levels relative to parental expression levels has been used. First, genes are divided into two groups based on the comparison of the relative expression level in the inbred parents. One group includes the set of genes that are expressed at equivalent levels in the two parents, while the second group includes the genes that are differentially expressed in the two parents. For genes expressed equally between inbred parents, there are three possible classifications of hybrid expression: (1) equal to the two parents (the gene would then be called "not differentially expressed" among the genotypes), (2) above the inbred expression levels (termed "above high-parent"), or (3) below the inbred expression levels (termed "below low-parent"). For genes that are differentially expressed between inbred parents, there are five possible classifications of hybrid expression: (1) above the higher inbred parent level (termed "above high-parent"), (2) similar to the higher inbred parent (termed "high-parent"), (3) in between the inbred parent levels (termed "between-parent"), (4) similar to the lower inbred parent level (termed "low-parent"), or (5) below the lower inbred parent level (termed "below low-parent").

Allelic variation may determine the mode of hybrid expression in a predictable manner. It is expected that *cis*-variant inbred alleles will maintain their expression differences following hybridization and produce additive patterns of gene expression in the hybrid, because of the averaging effect of having one higher-expressing allele and one lower-expression allele present. *Trans*-variant alleles can result in additive hybrid expression if the *trans*-regulator is dosage-dependent, or *trans*-variant alleles can result in nonadditive hybrid expression if the *trans*-regulator is dosage-independent. However, the sources of extreme hybrid expression patterns, such as above high-parent and below low-parent, are not understood. Such patterns could arise through epistatic interactions or through a combination of complementary *cis*- and *trans*-regulatory variation.

expression level is outside the range of the level of the parents. The first studies to provide evidence of this phenomenon focused on using 2D-PAGE to observe protein levels. Romagnoli et al. (1990) and Leonardi et al. (1991) used such methodologies to suggest that maize hybrids exhibit significant levels of nonadditive expression.

Hollick and Chandler (1998) provided evidence for overdominant action by some alleles at the *Pl* locus of maize. The anthocyanin levels were significantly higher in heterozygotes containing one copy of an epigenetically modified *Pl* allele than in either of the homozygous parent plants. The higher levels of pigmentation likely result from over-high-parent expression of the *Pl* gene in these hybrids relative to inbreds (Hollick and Chandler 1998).

Song and Messing (2003) focused on characterizing the relative expression of *zein1C* genes in developing endosperm tissue using an RT-PCR/sequencing approach. BAC sequencing was used to show that the B73 and BSSS53 genotypes encode a different set of *zein1C* genes. By performing reverse-transcriptase-mediated PCR and then sequencing large numbers of the resulting *zein1C* cDNAs, it was possible to determine which genes were expressed and their relative expression levels. B73 expresses six different *zein1C* genes, while BSSS53 expresses seven *zein1C* genes, but only three of these genes are shared between the genotypes. Analysis of these genes in hybrid endosperm tissue found that all 10 *zein1C* genes were expressed, but the relative expression levels were often nonadditive relative to the parental levels (Song and Messing 2003).

Auger et al. (2005a) used quantitative Northern blot hybridization to study gene expression levels in mature leaf tissue for ~30 genes, including nuclear-, mitochondrial-, and chloroplast-encoded genes. A surprisingly high level of nonadditive gene expression was documented in the reciprocal hybrids compared to their inbred parents, B73 and Mo17. More than three-quarters of the monitored genes (24/30) were found to display nonadditive expression patterns in at least one of the two reciprocal hybrids relative to the inbred parents. In addition, 16 of the 24 nonadditive genes displayed hybrid expression levels outside the range of the levels of the two parents (Auger et al. 2005a). This study suggested that maize hybrids may have different expression patterns from what would be predicted by the mid-parent values of the two inbred parents. Several specialized microarray approaches were used to document evidence for nonadditive gene expression in maize hybrids, although in most cases the nonadditive expression was within the range of the parental expression levels (Meyer et al. 2007; Użarowska et al. 2007).

A similar picture of prevalent nonadditive gene expression has been reported in microarray-based experiments designed to monitor the level of additive and nonadditive expression in *Arabidopsis*, *Drosophila*, and rice (Gibson et al. 2004; Ranz et al. 2004; Vuylsteke et al. 2005; Huang et al. 2006). A significant fraction (~10%) of the genes monitored in *Arabidopsis* displayed evidence for hybrid gene expression outside of the parental range (Vuylsteke et al. 2005). A well-designed experiment in *Drosophila melanogaster* also found prevalent nonadditive expression in hybrids. In fact, the gene expression patterns of the two inbred parents were more similar to each other's than to the hybrids' (Gibson et al. 2004). Analysis of the differentially expressed genes in this experiment found that the number of genes with above high-parent or below low-parent expression patterns was significantly higher than the number of genes with additive, high-parent or low-parent expression patterns. Almost one-third of the differentially expressed genes identified in rice displayed expression levels that deviated from the mid-parent levels (Huang et al. 2006). However, there was no evidence that the expression level for these genes was outside the range of the parental expression values. In combination, these experiments suggest that hybrids may have significant levels of nonadditive gene expression including many cases of expression outside of the parental range.

Evidence for prevalent additive expression in hybrids

In contrast to the above studies, several groups have reported that maize hybrids tend to exhibit frequent additive expression patterns. These studies do find evidence for nonadditive expression, but these are primarily examples of expression within the

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

range of the parents. Guo et al. (2003) used a cDNA-AFLP approach to analyze expression levels in maize endosperm tissue isolated from genotypes that display a range of heterosis for yield. The majority of genes were expressed at additive levels in the hybrids. By studying the reciprocal hybrids, they found that ~10% of genes are expressed at nonadditive levels; these genes typically displayed maternal-like or paternal-like expression levels. There was no significant correlation between the amount of heterosis and the frequency of nonadditive patterns (Guo et al. 2003). More recently, this technique was applied to profiling expression in immature ear tissue derived from a set of 16 hybrids (Guo et al. 2006). In this study, it was found that the majority of differentially expressed genes display hybrid expression patterns within the range of the parents. This study included a series of hybrids with a range of heterotic response, and it was demonstrated that the proportion of genes displaying mid-parent expression is positively correlated with heterosis and yield (Guo et al. 2006).

Two recent microarray studies have also found evidence for low levels of novel expression in hybrids relative to inbreds. Swanson-Wagner et al. (2006) used a cDNA microarray to monitor the expression levels of ~14,000 genes in 14-d seedlings of B73, Mo17, and the F_1 hybrid. Approximately 10% of the genes were identified as differentially expressed in at least one genotype. The majority of these differentially expressed genes (1062/1367) displayed an additive pattern of gene expression in the hybrids relative to the inbred parents. While most of the remaining genes exhibited expression patterns similar to high-parent or low-parent levels, there was a subset of 44 genes that displayed nonadditive expression patterns outside of the parental range. Most of the genes (42/44) that displayed above high-parent or below low-parent expression were cases in which the expression of B73 was not significantly different from Mo17 but the hybrid displayed a nonequivalent expression pattern. Within the set of genes that displayed novel hybrid expression levels, there were only a few examples with large fold changes (33 of 44 were <1.4-fold different from the near parent) (Swanson-Wagner et al. 2006).

Gene expression in 11-d seedlings, immature ear, and embryo tissue was monitored in essentially the same genotypes (B73, Mo17, and reciprocal F_1 hybrids) using Affymetrix microarrays (Stupar and Springer 2006). A series of statistical and clustering analyses indicated that most genes display additive expression patterns in the hybrids relative to the inbred parents. Nearly 80% of the genes that are differentially expressed in B73 relative to Mo17 demonstrate mid-parent expression in the hybrid. The vast majority of the genes that displayed nonadditive expression patterns were found to be expressed at levels within the range of the two parents. Additionally, this study found no evidence for novel hybrid expression patterns in genes that were not differentially expressed in the two inbred parents.

The results of these two microarray studies are similar in that the majority of genes that are differentially expressed in the inbred parents show an additive pattern of expression in the hybrids, and most of the genes with nonadditive expression are still expressed within the parental range. Another microarray experiment that compared maize hybrid and inbred expression levels using an Agilent platform supports the conclusion that novel hybrid expression patterns are rare (Ma et al. 2006).

A recent study from the mouse community also found evidence of prevalent additive expression levels in F_1 hybrids. Cui et al. (2006) used an Affymetrix mouse GeneChip platform to compare the gene expression of liver tissues from two relatively closely related mice inbred strains and their reciprocal F_1 hybrids. The authors found that the F_1 hybrids exhibited additive expression levels much more frequently than nonadditive levels. The few observations of hybrid gene expression outside of the parental range were typically identified in genes with little expression difference between the inbred parents (Cui et al. 2006), similar to the results obtained by Swanson-Wagner et al. (2006) in maize.

Epistasis in gene expression

While the concept of epistasis is relatively straightforward (see Text Box 1), the specific application to heterosis has been varied. We discuss here the potential of transcriptional epistasis, defined here as gene interactions that affect the transcription rate of another gene or set of genes. The anthocyanin biosynthesis pathway of maize provides an excellent example of transcriptional epistasis. The maize *B* and *Pl* gene products are transcription factors that interact to up-regulate the synthesis of genes such as *A1*, *A2*, and *Bz1*, which control anthocyanin production (Dooner et al. 1991). An inbred with a nonfunctional *b* allele would likely display a green phenotype due to the low or absent expression of *A1*, *A2*, and *Bz1*. Similarly, a second inbred fixed for a nonfunctional *pl* allele would also be green. A hybrid produced between these two plants would be *B/b Pl/pl* and have high levels of expression for *A1*, *A2*, and *Bz1* and a red pigmentation (Fig. 3). The expression phenotype of *A1*, *A2*, and *Bz1* would be above high-parent in the hybrid relative to the inbred parents, while the expression phenotype of *B* and *Pl* would be additive. This overdominance would be the result of epistatic interactions between *B* and *Pl*, not the result of direct allelic interactions at *A1*, *A2*, or *Bz1*.

One might expect to observe numerous instances of transcriptional epistasis in maize hybrids; the merger of two genomes may result in new interactions within gene regulation pathways that were not present in the inbred parents, as in the anthocyanin example described above. If such cases were common in hybrids, nonadditive expression profiles outside of the parental range would be frequently observed, as was the case in *Drosophila* (Gibson et al. 2004). Therefore, it was surprising that there are few examples of this type in the microarray studies of maize.

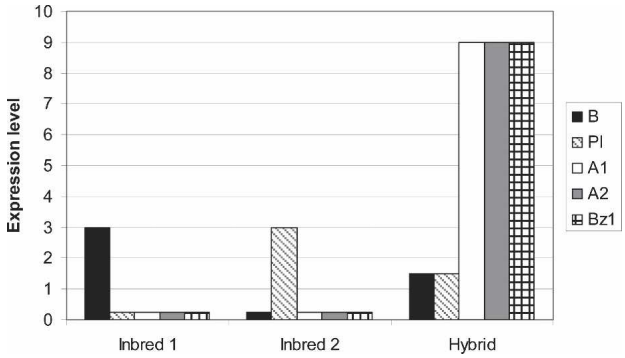

Figure 3. Example of potential transcriptional epistasis in maize hybrids. A potential example of transcriptional epistasis in maize is illustrated using genes from the well-characterized anthocyanin biosynthesis pathway. Inbred 1 (genotype *B*, *pl*, *A1*, *A2*, *Bz1*) and Inbred 2 (genotype *b*, *Pl*, *A1*, *A2*, *Bz1*) each contains a nonfunctional allele for one of the regulatory genes *B* or *Pl*, and both show negligible levels of expression for the structural genes *A1*, *A2*, and *Bz1*. The hybrid shows additive expression for *B* and *Pl* but nonadditive, over-high-parent expression for *A1*, *A2*, and *Bz1*.

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

Therefore, transcriptional epistasis, in which novel gene interactions affect the expression levels of downstream genes, is not common in hybrids of B73 and Mo17. Additionally, quantitative genetics analyses have suggested that epistasis does not play a major role in contributing to hybrid maize yield (Hinze and Lamkey 2003; Mihaljevic et al. 2005).

Despite multiple studies addressing the topic, a consensus view of gene expression in maize hybrids has not emerged. The prevalence of novel expression patterns in maize hybrids remains an open question. It is possible that the tissue sampling and experimental design differences have influenced the expression profiles observed in the studies discussed above.

Contribution of allelic variation to heterotic phenotypes

In the previous two sections, we have reviewed intraspecific variation in maize (allelic variation) and the relative expression patterns in inbreds relative to hybrids (allelic interactions). The subsequent question is to determine how allelic variation and interactions result in a superior hybrid phenotype. While outside the scope of this review, a complete model of heterosis undoubtedly involves contributions from quantitative genetics, physiology, and biochemistry.

Considerations of heterosis should take into account the following four factors.

(1) The magnitude of heterosis varies in different species. For example, heterosis has much stronger and more ubiquitous effects in maize than in *Arabidopsis* (Zanoni and Dudley 1989; Meyer et al. 2004).
(2) The level of heterosis for specific traits varies and is not correlated in different hybrids of the same species. This indicates that heterosis is not the result of action at a single locus, nor does heterosis simply reflect the overall extent of heterozygosity between parents.
(3) As the genetic distance between the parental inbreds increases, there is generally an increase in heterosis; yet when the parental distance exceeds a threshold, heterosis decreases (Moll et al. 1965). Thus, there appears to be a relationship between genetic diversity and heterosis; however, the correlation is not strong enough to be used as an accurate predictive tool (Melchinger 1999).
(4) The allelic variation that produces heterosis is not representative of all variation that occurs. Not all allelic variants that arise in a maize population will become fixed in inbred lines. Variants with strong deleterious phenotypes will be selected against by breeders. Therefore, the range of allelic variation that is captured in inbred lines that can affect heterosis is limited to the variation that has beneficial affects for a specific trait or has limited deleterious affects (Troyer 2006).

Collectively, these factors provide insight into several important features of heterosis. First, heterosis is the result of variation at multiple genomic locations. Second, many of the complex phenotypes that are often assessed for heterosis, such as yield, are likely influenced by many (hundreds of) genes. Third, heterosis for different phenotypes is determined by variation at a partially nonredundant set of loci. These features of heterosis are supported by QTL studies that have mapped QTLs controlling heterosis for yield, height, and flowering date in maize and rice (Stuber et al. 1992; Xiao et al. 1995; Li et al. 2001; Luo et al. 2001; Lu et al. 2003). Although there was evidence for overdominance in some of these studies, it is quite possible that this apparent overdominance could reflect pseudo-overdominance resulting from repulsion-phase linkages (see Text Box 1). A QTL fine-mapping approach confirmed that this was the case for one apparent overdominant locus (Graham et al. 1997).

Complementation of present–absent genes

The dominance hypothesis proposes that heterosis is the result of complementation of slightly deleterious alleles in the hybrid. The inbreds become fixed for these alleles, resulting in inbreeding depression. Fu and Dooner (2002) refined this model to include the potential of present–absent genes. There is growing evidence that the suite of active genes in heterotic hybrids is greater than that of either of the parental inbred lines.

Nonshared genic fragments in maize inbred lines (for reviews, see Buckler et al. 2006; Messing and Dooner 2006) provide instances of variation in which one inbred contains genic information that is absent in another inbred. The majority of these nonshared genic sequences appear to be novel sequence gain (duplication events) in one inbred relative to the other, as opposed to deletion events of previously shared genes (Lai et al. 2005). Novel functions may arise as gene fragments, or full-length genes, are copied to new chromosomal locations and potentially acquire novel expression patterns. In addition, the mechanism of transposition appears to have the capacity to perform exon shuffling and create new ORFs at a relatively high frequency (Brunner et al. 2005b; Lai et al. 2005).

In addition to the differences in genic content of the inbred genomes, there is also evidence for a significant number of genes that are expressed in one inbred line but not detected in another inbred (Song and Messing 2003; Stupar and Springer 2006). These genes with presence–absence inbred expression patterns are expressed at detectable levels in the hybrid. Both gain of transcripts (through altered regulation or new genic fragments) and loss of transcripts (through altered regulation or transposon insertions) would result in the presence–absence transcriptome differences between maize inbreds.

What are the specific mechanisms driving these differences in genomic and transcriptome content? There is evidence that a primary source of the variation in genome content, and potentially transcriptome content, is the action of transposable elements. *Helitron* elements, DNA transposons, and retrotransposons constitute two-thirds of the maize genome and are likely to be continuously generating variation. The "capture" and movement of genic fragments by DNA transposons, retrotransposons, and rolling circle transposons can increase the copy number for genic fragments (Bennetzen 2005). Furthermore, transposable element insertion events may cause a reduction in the expression level of genes through disruption of coding regions or regulatory regions. Alternatively, transposable elements may have more complex effects on gene regulation through the provision of novel regulatory sequences (Barkan and Martienssen 1991).

Complementation of novel beneficial alleles

One clue to the underlying causes of heterosis comes from the analysis of diversity limitations in heterotic crosses. Crosses between highly diverse maize germplasm often exhibit limited heterosis. Tropical germplasm has not been adapted to environmental conditions in the temperate United States, and this may cause

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

the inferior heterotic performance of these stocks. Troyer (2006) points out the importance of adapted, or beneficial, alleles in heterosis. Any breeding program will include artificial selection for specific traits in addition to inadvertent selection dictated by environmental factors. During the domestication and spread of maize, different inbreds have evolved unique sets of adaptations specific to the selective pressures of growth and reproduction in different habitats. These adaptations might include altered flowering time, disease resistance, growth habits, or stress tolerances. In some cases, different sets of genes may have even been selected for the same traits. If we assume that these traits behave in a dominant or codominant fashion, then the improved phenotype of hybrids may be the result of an increase in the cumulative adaptedness of hybrids for specific traits. The hybrid will contain one complete set of alleles from each of the inbred parents, resulting in a combination of each of the beneficial alleles accrued by each parent. This concept is simply a minor revision of the dominance model. In the dominance model, hybridization essentially masks the accumulated deleterious alleles in the inbreds. In the beneficial allele model, hybridization merges the accumulated specific and divergent beneficial alleles from both inbreds. This model could also be applied to overdominance. If two different alleles of the same gene acquire unique beneficial adaptations, the heterozygote (containing both adaptations) may be more beneficial than either of the parental alleles alone.

A potential example is provided by a study of phosphate acquisition efficiency in *Arabidopsis* (Narang and Altmann 2001). The F_1 of a cross of inbred ecotypes *C24* and *Col-0* displays heterosis for phosphate acquisition efficiency. *Col-0* and *C24* each have adaptations to improve nutrient uptake characteristics: longer overall root length for *Col-0* and longer root hair length plus higher phosphate transporter expression for *C24*. The combined inheritance of these three dominant physiological adaptations results in heterosis for phosphate acquisition efficiency.

Potential advantages of mid-parent expression levels

As discussed above, there is a surprisingly high level of variation in the content and structure of the repetitive regions between maize genes. The repetitive regions may generate metastable epigenetic or chromatin modifications. There is evidence that these regions can influence the expression of adjacent genes (Stam et al. 2002; Clark et al. 2006) and may explain the high rate of *cis*-linked regulatory variation observed in maize (Guo et al. 2004; Stupar and Springer 2006). The high frequency of maize *cis*-acting regulatory variation may, in turn, lead to another potential mechanism of heterosis, mid-parent levels of gene expression.

There is ample genetic evidence, such as the phenotypic alterations displayed by over- or underexpressed transgenes, that altered gene expression levels can affect phenotypes in a quantitative manner. For example, many of the genes involved in disease resistance or pathogen defense display altered fitness in response to very high or low levels of expression (Heidel et al. 2004). While altered expression can affect phenotype, there is also evidence that expression level variation within an acceptable range does not affect phenotype. The prevalence of recessive mutations relative to dominant or codominant mutations suggests that many genes have a range of acceptable gene expression levels in which equally good phenotypes are produced. It is only when expression varies above or below thresholds that phenotypic consequences will ensue.

We propose a model that includes the following assumptions: (1) Gene expression levels will show a high rate of change within a maize population; (2) the cause of many of these gene expression differences is *cis*-linked; (3) these expression level changes can affect the phenotype when they deviate outside of an optimal range of expression and may result in a minor loss of fitness; (4) a subset of genes that display excessively high or low expression levels will become fixed within different inbred lines; and (5) the combination of two alleles in a hybrid that vary because of *cis*-linked changes will result in mid-parent expression levels. Given these assumptions, a model can be developed that would predict increased fitness of the hybrid relative to one or both of the inbreds (Fig. 4). Inbred gene expression levels may frequently be altered and then subsequently fixed above or below the optimal range, which may result in subtle effects on the phenotype. Mid-parent expression levels in the hybrid may dilute the detrimental effects of over- or underexpression of these genes by causing the total expression level for a gene to be within the optimal range. The hybrid may exhibit dominant gain (genes A or B) or overdominant gain (gene C) due to the hybrid mid-parent expression level. This model proposes that there will be instances in which the hybrid mid-parent expression level is advantageous relative to one (examples of genes A and B) or both (example of gene C) of the inbred parents.

Concluding remarks

There are several potential mechanisms through which genomic variation could combine to produce a heterotic phenotype. Each of these mechanisms could occur at a subset of genes, and the combination of effects will result in heterosis. In addition, there are other nonmutually exclusive mechanisms such as altered protein–protein interactions, novel epigenetics states, siRNAs, or altered hormone levels that can profoundly alter phenotypes (Phillips 1999; Birchler et al. 2003; Osborn et al. 2003). As additional genomic and genetic resources become available, it is anticipated that a better understanding of heterosis will emerge. Given the rate and types of allelic variation observed in maize inbreds, there is substantial evidence that allelic variants at a

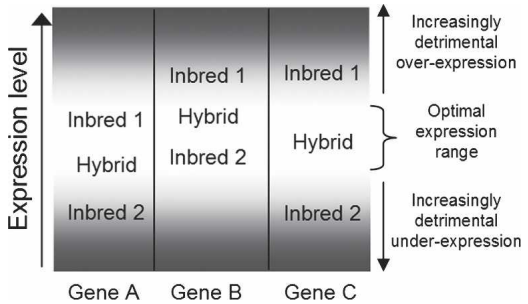

Figure 4. Potential model to explain how additive expression could be advantageous relative to one or both parents. Optimal gene expression levels (or patterns) occur in the white range at the *middle* of the graph. Increasing levels of shading indicate increasing deleterious effects on the phenotype due to low or high levels of expression. The exact range of optimal expression as well as the rate at which over- or underexpression becomes detrimental is likely to vary for different genes. We illustrate potential gene expression variation at three hypothetical genes. Gene A expression levels are within the optimal range for inbred 1 but under the optimal level for inbred 2. Gene B expression levels are within optimal range for inbred 2 but above the optimal level for inbred 1. Gene C exhibits suboptimal expression in both inbred 1 and inbred 2. The expression level in the hybrid is within the optimal range for all three genes.

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

large number of loci act through partial to complete dominance to provide favorable complementation resulting in superior hybrid phenotypes. The contributions of epistatic interactions and overdominance to heterosis are more difficult to establish and remain enigmatic.

Acknowledgments

We apologize to all those researchers whose work we have overlooked or could not include because of page limitations. This review was greatly improved by the comments of Virginia Walbot and two anonymous reviewers. We thank Shawn Kaeppler, William Tracy, Ronald Phillips, Ruth Shaw, Irina Makarevitch, Karen McGinnis, William Haun, Wade Odland, Peter Tiffin, and Cynthia Weinig for helpful comments and discussions on the manuscript. N.M.S. is supported by the National Science Foundation (grant no. DBI-0227310).

References

Adawy, S.S.M., Stupar, R.M., and Jiang, J. 2004. Fluorescence in situ hybridization analysis reveals multiple loci of knob-associated DNA elements in one-knob and knobless maize lines. *J. Histochem. Cytochem.* **52:** 1113–1116.

Auger, D.L., Gray, A.D., Ream, T.S., Kato, A., Coe Jr., E.H., and Birchler, J.A. 2005a. Nonadditive gene expression in diploid and triploid hybrids of maize. *Genetics* **169:** 389–397.

Auger, D.L., Peters, E.M., and Birchler, J.A. 2005b. A genetic test of bioactive gibberellins as regulators of heterosis in maize. *J. Hered.* **96:** 614–617.

Barkan, A. and Martienssen, R.A. 1991. Inactivation of maize transposon *Mu* suppresses a mutant phenotype by activating an outward-reading promoter near the end of *Mu1*. *Proc. Natl. Acad. Sci.* **88:** 3502–3506.

Bennetzen, J.L. 2005. Transposable elements, gene creation and genome rearrangement in flowering plants. *Curr. Opin. Genet. Dev.* **15:** 621–627.

Bingham, E.T. 1998. Role of chromosome blocks in heterosis and estimates of dominance and overdominance. In *Concepts and breeding of heterosis in crop plants* (eds. K.R. Lamkey and J.E. Staub), pp. 71–88. Crop Science Society of America, Madison, WI.

Birchler, J.A., Auger, D.L., and Riddle, N.C. 2003. In search of a molecular basis of heterosis. *Plant Cell* **15:** 2236–2239.

Bray, N.J., Buckland, P.R., Owen, M.J., and O'Donovan, M.C. 2003. *Cis*-acting variation in the expression of a high proportion of genes in human brain. *Hum. Genet.* **113:** 149–153.

Brown, W.L. 1949. Numbers and distribution of chromosome knobs in United States maize. *Genetics* **34:** 524–536.

Brunner, S., Fengler, K., Morgante, M., Tingey, S., and Rafalski, A. 2005a. Evolution of DNA sequence nonhomologies among maize inbreds. *Plant Cell* **17:** 343–360.

Brunner, S., Pea, G., and Rafalski, A. 2005b. Origins, genetic organization and transcription of a family of non-autonomous helitron elements in maize. *Plant J.* **43:** 799–810.

Buckler, E.S., Gaut, B.S., and McMullen, M.D. 2006. Molecular and functional diversity of maize. *Curr. Opin. Plant Biol.* **9:** 172–176.

Bureau, T.E., White, S.E., and Wessler, S.R. 1994. Transduction of a cellular gene by a plant retroelement. *Cell* **77:** 479–480.

Cheung, V.G. and Spielman, R.S. 2002. Genetics of variation in gene expression. *Nat. Genet.* **32 (Suppl):** 522–525.

Cheung, V.G., Spielman, R.S., Ewens, K.G., Weber, T.M., Morley, M., and Burdick, J.T. 2005. Mapping determinants of human gene expression by regional and genome-wide association. *Nature* **437:** 1365–1369.

Ching, A., Caldwell, K.S., Jung, M., Dolan, M., Smith, O.S., Tingey, S., Morgante, M., and Rafalski, A.J. 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet.* **3:** 19.

Clark, R.M., Wagler, T.N., Quijada, P., and Doebley, J. 2006. A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescent architecture. *Nat. Genet.* **38:** 594–597.

Coors, J.G. and Pandey, S. 1999. *Genetics and exploitation of heterosis in crops*. Crop Science Society of America, Madison, WI.

Cowles, C.R., Hirschhorn, J.N., Altshuler, D., and Lander, E.S. 2002. Detection of regulatory variation in mouse genes. *Nat. Genet.* **32:** 432–437.

Crow, J.F. 1999. Dominance and overdominance. In *The genetics and exploitation of heterosis in crops* (eds J.G. Coors and S. Pandey), pp. 49–58. Crop Science Society of America, Madison, WI.

Cui, X., Affourtit, J., Shockley, K.R., Woo, Y., and Churchill, G.A. 2006. Inheritance patterns of transcript levels in F_1 hybrid mice. *Genetics* **174:** 627–637.

Davenport, C.B. 1908. Degeneration, albinism and inbreeding. *Science* **28:** 454–455.

Dooner, H.K., Robbins, T.P., and Jorgensen, R.A. 1991. Genetic and developmental control of anthocyanin biosynthesis. *Annu. Rev. Genet.* **25:** 173–199.

Doss, S., Schadt, E.E., Drake, T.A., and Lusis, A.J. 2005. *Cis*-acting expression quantitative trait loci in mice. *Genome Res.* **15:** 681–691.

Duvick, D.N. 1999. Heterosis: Feeding people and protecting natural resources. In *The genetics and exploitation of heterosis in crops* (eds. J.G. Coors and S. Pandey), pp. 19–30. Crop Science Society of America, Madison, WI.

Duvick, D.N. 2001. Biotechnology in the 1930s: The development of hybrid maize. *Nat. Rev. Genet.* **2:** 69–74.

East, E.M. 1908. Inbreeding in corn. *Rep. Connecticut Agric. Exp. Stn.* **1907:** 419–429.

Emrich, S.J., Li, L., Wen, T.J., Yandeau-Nelson, M.D., Fu, Y., Guo, L., Chou, H.H., Aluru, S., Ashlock, D.A., and Schnable, P.S. 2007. Nearly identical paralogs (NIPs): implications for maize (*Zea mays* L.) genome evolution. *Genetics* doi: 10.1534/genetics.106.064006.

Flint-Garcia, S.A., Thuillet, A.C., Yu, J., Pressoir, G., Romero, S.M., Mitchell, S.E., Doebley, J., Kresovich, S., Goodman, M.M., and Buckler, E.S. 2005. Maize association population: A high-resolution platform for quantitative trait locus dissection. *Plant J.* **44:** 1054–1064.

Fu, H. and Dooner, H.K. 2002. Intraspecific violation of genetic colinearity and its implications in maize. *Proc. Natl. Acad. Sci.* **99:** 9573–9578.

Gibson, G., Riley-Berger, R., Harshman, L., Kopp, A., Vacha, S., Nuzhdin, S., and Wayne, M. 2004. Extensive sex-specific nonadditivity of gene expression in *Drosophila melanogaster*. *Genetics* **167:** 1791–1799.

Goldman, I.L. 1998. From out of old fields comes all this new corn: An historical perspective on heterosis in crop improvement. In *Concepts and breeding of heterosis in crop plants* (eds. K.R. Lamkey and J.E. Staub), pp. 1–13. Crop Science Society of America, Madison, WI.

Graham, G.I., Wolff, D.W., and Stuber, C.W. 1997. Characterization of a yield quantitative trait locus on chromosome 5 of maize by fine mapping. *Crop Sci.* **37:** 1601–1610.

Guo, M., Rupe, M.A., Danilevskaya, O.N., Yan, X., and Hu, Z. 2003. Genome-wide mRNA profiling reveals heterochronic allelic variation and a new imprinted gene in hybrid maize endosperm. *Plant J.* **36:** 30–44.

Guo, M., Rupe, M.A., Zinselmeier, C., Habben, J., Bowen, B.A., and Smith, O.S. 2004. Allelic variatioin of gene expressioin in maize hybrids. *Plant Cell* **16:** 1707–1716.

Guo, M., Rupe, M.A., Yang, X., Crasta, O., Zinselmeier, C., Smith, O.S., and Bowen, B. 2006. Genome-wide transcript analysis of maize hybrids: Allelic additive gene expression and yield heterosis. *Theor. Appl. Genet.* **113:** 831–845.

Hansson, B. and Westerberg, L. 2002. On the correlation between heterozygosity and fitness in natural populations. *Mol. Ecol.* **11:** 2467–2474.

Heidel, A.J., Clarke, J.D., Antonovics, J., and Dong, X. 2004. Fitness costs of mutations affecting the systemic acquired resistance pathway in *Arabidopsis thaliana*. *Genetics* **168:** 2197–2206.

Hinze, L.L. and Lamkey, K.R. 2003. Absence of epistasis for grain yield in elite maize hybrids. *Crop Sci.* **43:** 46–56.

Hollick, J.B. and Chandler, V.L. 1998. Epigenetic allelic states of a maize transcriptional regulatory locus exhibit overdominant gene action. *Genetics* **150:** 891–897.

Huang, Y., Zhang, L., Zhang, J., Yuan, D., Xu, C., Li, X., Zhou, D., Wang, S., and Zhang, Q. 2006. Heterosis and polymorphisms of gene expression in an elite rice hybrid as revealed by a microarray analysis of 9198 unique ESTs. *Plant Mol. Biol.* **62:** 579–591.

Janick, J. 1998. Hybrids in horticultural crops. In *Concepts and breeding of heterosis in crop plants* (eds. K.R. Lamkey and J.E. Staub), pp. 45–56. Crop Science Society of America, Madison, WI.

Jiang, N., Bao, Z., Zhang, X., Eddy, S.R., and Wessler, S.R. 2004. Pack-MULE transposable elements mediate gene evolution in plants. *Nature* **431:** 569–573.

Jin, Y.K. and Bennetzen, J.L. 1994. Integration and nonrandom mutation of a plasma membrane proton ATPase gene fragment within the *Bs1* retroelement of maize. *Plant Cell* **6:** 1177–1186.

Jones, D.F. 1917. Dominance of linked factors as a means of accounting

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

for heterosis. *Genetics* **2:** 223–238.
Kapitonov, V.V. and Jurka, J. 2001. Rolling-circle transposons in eukaryotes. *Proc. Natl. Acad. Sci.* **98:** 8714–8719.
Kato, A., Lamb, J.C., and Birchler, J.A. 2004. Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. *Proc. Natl. Acad. Sci.* **101:** 13554–13559.
Lai, J., Li, Y., Messing, J., and Dooner, H.K. 2005. Gene movement by *Helitron* transposons contributes to the haplotype variability of maize. *Proc. Natl. Acad. Sci.* **102:** 9068–9073.
Laurie, D.A. and Bennett, M.D. 1985. Nuclear DNA content in the genera *Zea* and *Sorghum*. Intergeneric, interspecific and intraspecific variation. *Heredity* **55:** 307–313.
Lee, J.H., Arumuganathan, K., Kaeppler, S.M., Park, S.W., Kim, K.Y., Chung, Y.S., Kim, D.H., and Fukui, K. 2002. Variability of chromosomal DNA contents in maize (*Zea mays L.*) inbred and hybrid lines. *Planta* **215:** 666–671.
Leonardi, A., Damerval, C., Herbert, Y., Gallais, A., and de Vienne, D. 1991. Association of protein amount polymorphisms (PAP) among maize lines with performances of their hybrids. *Theor. Appl. Genet.* **82:** 552–560.
Li, Z.K., Luo, L.J., Mei, H.W., Wang, D.L., Shu, Q.Y., Tabien, R., Zhong, D.B., Ying, C.S., Stansel, J.W., Khush, G.S., et al. 2001. Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. *Genetics* **158:** 1737–1753.
Lo, H.S., Wang, Z., Hu, Y., Yang, H.H., Gere, S., Buetow, K.H., and Lee, M.P. 2003. Allelic variation in gene expression is common in the human genome. *Genome Res.* **13:** 1855–1862.
Lu, H., Romero-Severson, J., and Bernardo, R. 2003. Genetic basis of heterosis explored by simple sequence repeat markers in a random-mated maize population. *Theor. Appl. Genet.* **107:** 494–502.
Luo, L.J., Li, Z.K., Mei, H.W., Shu, Q.Y., Tabien, R., Zhong, D.B., Ying, C.S., Stansel, J.W., Khush, G.S., and Paterson, A.H. 2001. Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components. *Genetics* **158:** 1755–1771.
Ma, J., Morrow, D.J., Fernandes, J., and Walbot, V. 2006. Comparative profiling of the sense and antisense transcriptome of maize lines. *Genome Biol.* **7:** R22.
McClintock, B., Kato, Y., and Blumenshein, A. 1981. *Chromosome constitution of races of maize*. Colegio de Postgraduados, Chapingo, Mexico.
Melchinger, A.E. 1999. Genetic diversity and heterosis. In *The genetics and exploitation of heterosis in crops* (eds. J.G. Coors and S. Pandey), pp. 99–118. Crop Science Society of America, Madison, WI.
Melchinger, A.E. and Gumber, R.K. 1998. Overview of heterosis and heterotic groups in agronomic crops. In *Concepts and breeding of heterosis in crop plants* (eds. K.R. Lamkey and J.E. Staub), pp. 29–44. Crop Science Society of America, Madison, WI.
Messing, J. and Dooner, H.K. 2006. Organization and variability of the maize genome. *Curr. Opin. Plant Biol.* **9:** 157–163.
Meyer, R.C., Torjek, O., Becher, M., and Altmann, T. 2004. Heterosis of biomass production in *Arabidopsis*. Establishment during early development. *Plant Physiol.* **134:** 1813–1823.
Meyer, S., Pospisil, H., and Scholten, S. 2007. Heterosis associated gene expression in maize embryos 6 days after fertilization exhibits additive, dominant and overdominant pattern. *Plant Mol. Biol.* doi: 10.1007/s11103-006-9095-x.
Mihaljevic, R., Utz, H.F., and Melchinger, A.E. 2005. No evidence for epistasis in hybrid and per se performance of elite European Flint maize inbreds from generation means and QTL analyses. *Crop Sci.* **45:** 2605–2613.
Mitton, J.B. 1998. Apparent overdominant in natural plant populations. In *Concepts and breeding of heterosis in crop plants* (eds. K.R. Lamkey and J.E. Staub), pp. 57–70. Crop Science Society of America, Madison, WI.
Mitton, J.B. and Jeffers, R.M. 1989. The genetic consequences of mass selection for growth rate in Engelmann spruce. *Silvae Genet.* **38:** 6–12.
Moll, R.H., Longquist, J.V., Fortuna, J.V., and Johnson, E.C. 1965. The relation of heterosis and genetic divergence in maize. *Genetics* **52:** 139–144.
Monks, S.A., Leonardson, A., Zhu, H., Cundiff, P., Pietrusiak, P., Edwards, S., Phillips, J.W., Sachs, A., and Schadt, E.E. 2004. Genetic inheritance of gene expression in human cell lines. *Am. J. Hum. Genet.* **75:** 1094–1105.
Mopper, S., Mitton, J.B., Whitham, T.G., Cobb, N.S., and Christensen, K.M. 1991. Genetic differentiation and heterozygosity in Pinyon pine associated with resistance to herbivory and environmental stress. *Evolution Int. J. Org. Evolution* **45:** 989–999.
Morgante, M., Brunner, S., Pea, G., Fengler, K., Zuccolo, A., and Rafalski, A. 2005. Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. *Nat. Genet.* **37:** 977–1002.
Morley, M., Molony, C.M., Weber, T.M., Devlin, J.L., Ewens, K.G., Spielman, R.S., and Cheung, V.G. 2004. Genetic analysis of genome-wide variation in human gene expression. *Nature* **430:** 743–747.
Narang, R.A. and Altmann, T. 2001. Phosphate acquisition heterosis in *Arabidopsis thaliana*: A morphological and physiological analysis. *Plant Soil* **234:** 91–97.
Osborn, T.C., Pires, J.C., Birchler, J.A., Auger, D.L., Chen, Z.J., Lee, H.S., Comai, L., Madlung, A., Doerge, R.W., Colot, V., et al. 2003. Understanding mechanisms of novel gene expression in polyploids. *Trends Genet.* **19:** 141–147.
Pastinen, T., Sladek, R., Gurd, S., Sammak, A., Ge, B., Lepage, P., Lavergne, K., Villeneuve, A., Gaudin, T., Brandstrom, H., et al. 2004. A survey of genetic and epigenetic variation affecting human gene expression. *Physiol. Genomics* **16:** 184–193.
Pastinen, T., Ge, B., Gurd, S., Gaudin, T., Dore, C., Lemire, M., Lepage, P., Harmsen, E., and Hudson, T.J. 2005. Mapping common regulatory variants to human haplotypes. *Hum. Mol. Genet.* **14:** 3963–3971.
Phillips, R.L. 1999. Research needs in heterosis. In *The genetics and exploitation of heterosis in crops* (eds. J.G. Coors and S. Pandey), pp. 501–508. Crop Science Society of America, Madison, WI.
Ranz, J.M., Namgyal, K., Gibson, G., and Hartl, D.L. 2004. Anomalies in the expression profile of interspecific hybrids of *Drosophila melanogaster* and *Drosophila simulans*. *Genome Res.* **14:** 373–379.
Romagnoli, S., Maddaloni, M., Livini, C., and Motto, M. 1990. Relationship between gene expression and hybrid vigor in primary root tips of young maize (*Zea mays* L.) plantlets. *Theor. Appl. Genet.* **80:** 769–775.
Schadt, E.E., Monks, S.A., Drake, T.A., Lusis, A.J., Che, N., Collnayo, V., Ruff, T.G., Milligan, S.B., Lamb, J.R., Cavet, G., et al. 2003. Genetics of gene expression surveyed in maize, mouse, and man. *Nature* **422:** 297–302.
Semel, Y., Nissenbaum, J., Menda, N., Zinder, M., Krieger, U., Issman, N., Pleban, T., Lippman, Z., Gur, A., and Zamir, D. 2006. Overdominant quantitative trait loci for yield and fitness in tomato. *Proc. Natl. Acad. Sci.* **103:** 12981–12986.
Shull, G.H. 1908. The composition of a field of maize. *American Breeders Assoc. Rep.* **4:** 296–301.
Song, R. and Messing, J. 2003. Gene expression of a gene family in maize based on non-collinear haplotypes. *Proc. Natl. Acad. Sci.* **100:** 9055–9060.
Stam, M., Belele, C., Dorweiler, J.E., and Chandler, V.L. 2002. Differential chromatin structure within a tandem array 100 kb upstream of the maize *b1* locus is associated with paramutation. *Genes & Dev.* **16:** 1906–1918.
Stranger, B.E., Forrest, M.S., Clark, A.G., Minichiello, M.J., Deutsch, S., Lyle, R., Hunt, S., Kahl, B., Antonarakis, S.E., Tavare, S., et al. 2005. Genome-wide associations of gene expression variation in humans. *PLoS Genet.* **1:** e78.
Stuber, C.W., Lincoln, S.E., Wolff, D.W., Helentjaris, T., and Lander, E.S. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics* **132:** 823–839.
Stupar, R.M. and Springer, N.M. 2006. *Cis*-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F_1 hybrid. *Genetics* **173:** 2199–2210.
Swanson-Wagner, R.A., Jia, Y., DeCook, R., Borsuk, L.A., Nettleton, D., and Schnable, P.S. 2006. All possible modes of gene action are observed in a global comparison of gene expression in a maize F_1 hybrid and its inbred parents. *Proc. Natl. Acad. Sci.* **103:** 6805–6810.
Tenaillon, M.I., Sawkins, M.C., Long, A.D., Gaut, R.L., Doebley, J.F., and Gaut, B.S. 2001. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proc. Natl. Acad. Sci.* **98:** 9161–9166.
Troyer, A.F. 2006. Adaptedness and heterosis in corn and mule hybrids. *Crop Sci.* **46:** 528–543.
Użarowska, A., Keller, B., Piepho, H.P., Schwarz, G., Ingvardsen, C., Wenzel, G., and Lubberstedt, T. 2007. Comparative expression profiling in meristems of inbred-hybrid triplets of maize based on morphological investigations of heterosis for plant height. *Plant Mol. Biol.* **63:** 21–34.
Vroh Bi, I., McMullen, M.D., Sanchez-Villeda, H., Schroeder, S., Gardiner, J., Polacco, M., Soderlund, C., Wing, R., Fang, Z., and Coe Jr., E.H. 2005. Single nucleotide polymorphisms and insertion–deletions for genetic markers and anchoring the maize fingerprint contig physical map. *Crop Sci.* **46:** 12–21.
Vuylsteke, M., van Eeuwijk, F., Van Hummelen, P., Kuiper, M., and

11/13

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

Zabeau, M. 2005. Genetic analysis of variation in gene expression in *Arabidopsis thaliana*. *Genetics* **171:** 1267–1275.
Wang, Q. and Dooner, H.K. 2006. Remarkable variation in maize genome structure inferred from haplotype diversity at the bz locus. *Proc. Natl. Acad. Sci.* **103:** 17644–17649.
Wittkopp, P.J., Haerum, B.K., and Clark, A.G. 2004. Evolutionary changes in *cis* and *trans* gene regulation. *Nature* **430:** 85–88.
Wright, S.I., Bi, I.V., Schroeder, S.G., Yamasaki, M., Doebley, J.F., McMullen, M.D., and Gaut, B.S. 2005. The effects of artificial selection on the maize genome. *Science* **308:** 1310–1314.
Xiao, J., Li, J., Yuan, L., and Tanksley, S.D. 1995. Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. *Genetics* **140:** 745–754.
Yan, H., Yuan, W., Velculescu, V.E., Vogelstein, B., and Kinzler, K.W. 2002. Allelic variation in human gene expression. *Science* **297:** 1143.
Zanoni, U. and Dudley, J.W. 1989. Comparison of different methods of identifying inbreds useful for improving elite maize hybrids. *Crop Sci.* **29:** 577–582.
Zhao, W., Canaran, P., Jurkuta, R., Fulton, T., Glaubitz, J., Buckler, E., Doebley, J., Gaut, B., Goodman, M., Holland, J., et al. 2006. Panzea: A database and resource for molecular and functional diversity in the maize genome. *Nucleic Acids Res.* **34:** D752–D757.

Downloaded from genome.cshlp.org on June 8, 2024 - Published by Cold Spring Harbor Laboratory Press

GENOME RESEARCH

Allelic variation and heterosis in maize: How do two halves make more than a whole?

Nathan M. Springer and Robert M. Stupar

Genome Res. 2007 17: 264-275 originally published online January 25, 2007
Access the most recent version at doi:10.1101/gr.5347007

References This article cites 72 articles, 35 of which can be accessed free at:
http://genome.cshlp.org/content/17/3/264.full.html#ref-list-1

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

To subscribe to *Genome Research* go to:
https://genome.cshlp.org/subscriptions

Copyright © 2007, Cold Spring Harbor Laboratory Press