Review Article

Heterosis: Many Genes, Many Mechanisms—End the Search for an Undiscovered Unifying Theory

Shawn Kaeppler

Department of Agronomy, University of Wisconsin, 1575 Linden Drive, Madison, WI 53706, USA

Correspondence should be addressed to Shawn Kaeppler, smkaeppl@wisc.edu

Heterosis is the increase in vigor that is observed in progenies of matings of diverse individuals from different species, isolated populations, or selected strains within species or populations. Heterosis has been of immense economic value in agriculture and has important implications regarding the fitness and fecundity of individuals in natural populations. Genetic models based on complementation of deleterious alleles, especially in the context of linkage and epistasis, are consistent with many observed manifestations of heterosis. The search for the genes and alleles that underlie heterosis, as well as for broader allele-independent, genomewide mechanisms, has encompassed many species and systems. Common themes across these studies indicate that sequence diversity is necessary but not sufficient to produce heterotic phenotypes, and that the molecular pathways that produce heterosis involve chromatin modification, transcriptional control, translation and protein processing, and interactions between and within developmental and biochemical pathways. Taken together, there are many and diverse molecular mechanisms that translate DNA into phenotype, and it is the combination of all these mechanisms across many genes that produce heterosis in complex traits.

1. Introduction

Heterosis has been observed and, in some cases, harnessed in many diverse systems. Examples of interspecies crosses of mammals that produce heterotic phenotypes include the mule resulting from a cross between a male donkey and a female horse, and the liger resulting from a cross between a lion and a tiger. In both cases, these interspecific hybrids are larger and, by some measures, more vigorous than the parents. However, many interspecific hybrids suffer from reduced longevity and reductions in fertility. Heterosis in humans has been proposed, sometimes controversially, to affect multiple phenotypes including attractiveness [1], IQ [2, 3], and height [4–6]. In agricultural settings, there are numerous examples in which heterosis has been harnessed to create more productive and more uniform products including livestock [7–11] and crop plants (reviewed in [12–19]). Heterosis can also be captured and fixed through the process of polyploidization which is common in the plant kingdom (reviewed in [13, 14, 20]). In this case, hybrids formed by sexual combination of unreduced gametes or by hybridization followed by chromosome doubling are often fertile and have often been classified as a new species. Polyploid individuals show a general trend toward an increase in size, and the capture of heterotic genetic effects can further enhance their fitness and productivity.

The impressive phenotypic manifestations of heterotic hybrids coupled with the economic importance of hybrid strains have led to extensive research to understand its basis. This research has followed evolving knowledge of genome composition and genetic and biochemical mechanisms and is enabled by technical advances that facilitate new measurements of phenotypes and molecular processes.

2. How Is Heterosis Defined?

Historical accounts of the development of the modern concept of heterosis are provided in several excellent articles [15, 21–23]. Documentation of the importance of inbreeding and performance included descriptions by early agriculturists who noted the deleterious effects of inbreeding in both plants and animals and took measures to minimize

this effect. Collins [24] documents activities of primitive tribes to mitigate inbreeding and maximize heterosis by placing seeds of multiple strains within each hill of maize that they planted. Darwin [25] experimentally evaluated the detrimental effect of mating among relatives supporting the idea that genetic diversity is related to hybrid vigor. Research on maize was important in developing some of the early ideas of heterosis [26–29] and has continued as an important experimental organism to the present. Shull's [30] article "The composition of a field of maize" is widely regarded as it provided seminal foundational ideas for inbreeding and hybridization in crop plants and was important in nucleating research in the recent era.

The fundamental concept of heterosis, as envisioned by Shull, is that deleterious alleles persist in large random-mating populations. Inbreeding due to drift, population isolation, or consanguineous mating by plan or by chance reduces vigor of individuals or populations due to increasing homozygosity of deleterious alleles. Vigor is restored by crossing among divergent types as recessive deleterious alleles are complemented in the hybrid state. This fundamental idea is consistent with many examples of heterosis across species.

Heterosis is quantified on an individual or population basis as the difference in the performance of the hybrid relative to the average of the inbred parents (termed the mid-parent value). For quantitative genetic analysis, the deviation of the hybrid relative to the mid-parent is the relevant value. In a practical context, high-parent heterosis, which measures the superiority of the hybrid relative to the best parent, is the important metric.

The conceptual opposite of heterosis is inbreeding depression [31]. This is the loss of vigor following related matings. Heterosis is often viewed as maximizing heterozygosity and, in contrast, inbreeding depression is due to reduction in heterozygosity. Inbreeding depression is measured as the reduction in performance in proportion to reduction in heterozygosity. Inbreeding depression is important in many settings including agriculture such as in maintenance of heirloom varieties, conservation biology, and human health. In any circumstance in which matings occur in small populations and/or assortative mating occurs, there is an increased risk of reduction in vigor and homozygosity of deleterious alleles in genotypic contexts that are otherwise rare in populations.

It is important to emphasize that measures of heterosis are phenotype-dependent. For example, interspecific mammalian hybrids may display increased size, vigor, and other desirable fitness traits, but have are often highly sterile and therefore have reduced fecundity. Flint-Garcia et al. [32] measured 17 traits among 267 maize hybrids and found that the amount of heterosis in any hybrid relative to its parents was trait-dependent and that hybrids could not be simply classified as heterotic or nonheterotic. From a research standpoint, this indicates that the search for mechanisms of heterosis must be conducted within the biological context of specific traits; in a practical context, it motivates research to better predict heterotic hybrids that will provide maximum productivity for specific traits of interest.

3. A Case for the Dominance Hypothesis: Early 1900s to Present

According to quantitative genetic theory, heterosis can result from dominance, overdominance, or epistasis. Overdominance is an intra-allelic interaction in which the presence of multiple alleles leads to greater performance than homozygosity for either allelic state. If overdominance is the predominant basis of heterosis, then populations and breeding strategies that maximize heterozygosity will result in the best performance. On the other hand, if dominance or epistasis is the primary mechanism of heterosis, natural or breeding populations, and therefore individuals, will become fixed for favorable alleles and perform equally to any hybrid. This issue was addressed from the early to mid-1900s by analysis of variance components (summarized in Hallauer et al. [33]).

Variance decomposition studies in hybrid maize populations using mating designs such as the North Carolina Design III resulted in significant estimates of overdominant gene action (summarized in [33]). However, Moll et al. [34] and Gardner and Lonnquist [35] realized that variance estimates could be confounded by linkage. Specifically, if positive and negative alleles were in repulsion phase linkage and the gene action of each locus was partial to complete dominance, the alleles at the two loci would frequently segregate together resulting in estimates of overdominance. In the Moll et al. [34] and Gardner and Lonnquist [35] studies, the average degree of dominance was estimated in the first generation of intermating from a population cross and then after intermating incrementally for multiple generations. The result of these studies was that the estimate of the average degree of dominance decreased, consistent with partial dominance—not overdominance—of most loci contributing to heterosis coupled with repulsion phase linkage.

The importance of dominance versus overdominance was further supported by recurrent selection studies in which populations were evaluated in crosses with each other, or with an inbred tester. Research by Russell et al. [36] in maize supported dominance versus overdominance as the primary basis of heterosis. One component of their study was the comparison of response to selection of populations selected based on performance of a cross with an inbred tester versus a population tester. If overdominance is the primary mechanism of heterosis, then the inbred tester would improve the population more than the population tester because in an inbred, alleles are fixed whereas in a population, they are intermediate in frequency. The result of this component of the study was that the inbred and population tester improved performance of the population similarly, consistent with the importance of dominance relative to overdominance. A second component of the Russell et al. [36] study was the analysis of selection in two populations based on performance of the population cross. If overdominance was the primary basis of heterosis, the populations would diverge due to selection and increase homozygosity of alternative alleles within the populations to maximize heterozygosity and performance of the population cross. The

result would be increasing performance of the population cross and decreasing performance of the populations *per se*. Alternatively, if dominance (or epistasis) were the primary mechanism of heterosis, the frequency of the favorable allele would increase in each population, and therefore also in the population cross, resulting in increasing performance of the populations and populations cross. The result of their study found increasing performance in all populations, supporting the importance of dominance versus overdominance. Note also that the level of linkage disequilibrium in these materials was likely quite low, minimizing confounding effects of pseudooverdominance.

Quantitative trait locus (QTL) mapping studies in maize are also consistent with dominance versus overdominance as the prevalent type of gene action underlying heterosis for productivity. Initial QTL studies indicated many QTL with overdominant gene action in populations derived from heterotic maize hybrids for traits such as yield and plant height [37, 38]. However, subsequent genetic dissection of an QTL with estimated overdominant gene action showed that the original QTL could be separated into two, linked QTLs in repulsion phase with dominant gene action [39, 40] conducted a QTL mapping study using 3 recombinant inbred populations using a North Carolina Design III approach. The results of this study were consistent with previous studies in maize. Overdominant gene action was estimated for QTL controlling grain yield, but those QTL were found in centromeric regions with high linkage disequilibrium (LD) and were interpreted as pseudooverdominance. Consistent with many other studies, the degree of heterosis was trait-dependent, with greatest heterosis for yield. Therefore, recent QTL mapping studies in maize are also generally consistent with a prevalence of dominance underlying heterotic traits including yield and yield components and growth traits such as plant height.

Xiao et al. [41] evaluated heterosis for ten traits *per se* using a testcross evaluation of a recombinant inbred line population derived from an interspecific *indica* × *japonica* cross in rice. The authors concluded that dominance was the primary basis of heterosis in this cross based on evidence from QTL, the absence of significant digenic epistatic interactions, and the relatively low relationship between marker heterozygosity and performance for most traits. Furthermore, two inbred lines from the population exceeded the performance of the hybrid, consistent with the proposition that, under the dominance hypothesis, it is possible to produce a homozygous individual that contains all the favorable alleles that produced the observed hybrid performance.

Despite a preponderance of evidence for the role of dominance in heterosis for yield in plants, especially in the context of linkage resulting in pseudooverdominance, there are observations that are inconsistent with the dominance hypothesis. One important observation is that, in some hybrids, the performance of the hybrid is greater than the sum of the parents. Given complete dominance, the maximum performance of the hybrid would be equal to the sum of the parents. Furthermore, as described below, well-documented examples of overdominance exist, and there is growing evidence for the role of epistasis using new experimental and statistical approaches.

4. Overdominance: Rationale and Examples

Overdominance is conceptually consistent with the idea that genetic dissimilarity *per se* stimulates vigor and, in a practical context, the optimum genetic state is heterozygosity versus homozygosity for favorable alleles. Overdominance provides an explanation for examples in which hybrid performance is greater than the sum of the parents, an incongruity with the dominance hypothesis.

Estimates of overdominant gene action have now generally been attributed to pseudooverdominance as described above. However, intriguing examples of overdominance have been reported. A biochemical example of overdominance provided by Schwartz and Laughner [42] was intellectually important in fueling the ongoing debate regarding the basis of heterosis. This study involved the activity of the enzyme *adh1*, which functions as a heterodimer. An allele of the enzyme with high activity was combined with an allele that had heat tolerance. The activity of the resulting biallelic enzyme was superior to that of either monoallelic form under specific stress conditions. This result provides a conceptual basis to consider molecular mechanisms by which intra-allelic interactions would provide increased performance and stress tolerance.

Krieger et al. [43] reported a single-gene model for overdominance based on developmental timing. In this study, heterozygosity for a functional allele and a loss-of-function allele at the *single flower truss* (*SFT*) locus in tomato results in overdominant fruit yield. This gene is homologous to *Arabidopsis Flowering Locus T* (*FT*) which is involved in the production of the flowering hormone florigen. Overdominant gene action for yield, in this example, is a result of shifting the developmental program so that an increased number of flowering inflorescences can form in the heterozygote relative to the wild-type homozygote which ends inflorescence production earlier and the mutant homozygote that produces limited inflorescences and more vegetative growth. In contrast to the specific example of an intra-allelic interaction in the case of *adh1* heterosis, the *SFT* result is based on dosage-dependent molecular expression (possibly additive) that results in a balance of gene product that is manifested in an overdominant phenotype. The *SFT* result is also compelling as there are likely multiple examples of intra- and inter-specific hybrids in which loss-of-function or allelic absence due to presence/absence variation (PAV) are combined in hybrids with a functional allele. Finally, this example highlights the potential productivity outcomes of fine-tuning developmental programs.

Semel et al. [44] evaluated gene action for 35 traits in tomato using an introgression line population in which each line of the cultivated tomato (*Solanum lycopersicum*) parent contained a small contribution from the genome of the wild species *Solanum pennellii*. The introgression lines were crossed to a cultivated line to produce hybrids. Most of the reproductive traits related to seed and fruit

yield exhibited overdominance, while nonreproductive traits related primarily to morphological characteristics did not. Based on the fact that some traits exhibited overdominance while others did not, the authors argued that this study supported true overdominance as opposed to pseudooverdominance. Additional research is required to assess whether this interpretation is correct.

These examples and others not included here provide evidence that overdominance can play a role in heterosis. However, the majority of the studies to date, based on response to selection, genetic variance partitioning, and QTL mapping are consistent with a lesser role for overdominance than dominance.

5. Epistasis: Emerging Evidence for the Role of Epistasis in Heterosis

The role of epistasis in heterosis remains elusive, although recent experiments provide increasing evidence for its importance. Estimates of epistatic variance in early studies of heterosis were limited by experiment size and computational capacity. Recent studies utilizing molecular markers and modern, computationally intensive statistical approaches, have increased ability to detect epistatic interactions.

Generation means analysis provided some of the first compelling evidence for the role of epistasis in hybrid performance. A recent example provided by Wolf and Hallauer [45] used a means-based analysis to support of role of epistasis in heterosis. The triple testcross analysis compares the relative performance of segregating progeny when testcrossed to both parents and to the F_1 hybrid. Deviation in performance of the F_1 testcross from the average of the parental testcrosses is consistent with epistatic gene action. Using this approach, the authors detected epistasis for multiple traits including yield, yield components, and timing of development among progeny of the heterotic hybrid B73 × Mo17.

Recent studies in maize, rice, and *Arabidopsis* based on QTL mapping report epistasis for various traits. Kusterer et al. [46] used a triple-testcross design in the context of QTL analysis in Arabidopsis to characterize the importance of epistasis for biomass traits. This research was complemented by a related study of near isogenic lines [47, 48]. Recent QTL mapping studies support the role of epistasis in rice [49–51]. The type of epistasis varies in these studies, from primarily additive × additive epistasis to dominant epistatic interactions, at least in part due to experimental materials and approach. Yu et al. [49] evaluated inbred F_2-derived F_3 families from the intraspecific cross Zhenshan97 × Minghui63 and reported a predominance of additive × additive interactions underlying performance for grain yield. In contrast, Li et al. [50] evaluated backcross (BC) and testcross hybrids from progeny of an interspecific *japonica* × *indica* hybrid and reported overdominant epistatic interactions. Hua et al. [51] evaluated an "immortalized F_2" population based on intermating recombinant-inbred lines and reported the important role in dominant × dominant epistatic interactions. Interpreting and summarizing trends across these studies, (1) interspecific populations whose parents have been genetically separated for a greater period of time exhibit more segregation and a greater degree of epistatic gene action, (2) experimental designs which utilize individuals with more heterozygosity (testcross or intermated RIL) detect higher levels of dominance, and (3) interpretation of overdominance remains confounded with pseudooverdominance in most studies.

It is logical to consider the potential relevance in the context of metabolic and physiological pathways. One physiological pathway that has been studied specifically in the context of heterosis is gibberellic acid (GA) metabolism and signaling. Production of GA involves a multistep pathway, and transduction of the GA signal involves a complex signaling network. Therefore, this metabolic and signaling pathway provides ample opportunity for the expression of epistatic gene action. In maize, inbreds contain less endogenous GA and precursors than corresponding hybrids [52]. Application of exogenous GA stimulates growth of inbreds more than hybrids [53, 54], consistent with the hypothesis that the reduced efficiency of inbreds to produce GA results in reduced biomass accumulation. A recent study in rice provides similar support for the role of GA in heterosis for biomass accumulation [55]. This study provided metabolic and transcriptome evidence to support the importance of GA synthesis and signaling in heterosis during rice seedling development.

The role of epistasis in heterotic and nonheterotic trait performance remains intriguing and perplexing. Conceptually, it is clear that many and diverse complex pathways interact to produce phenotypes in individuals supporting the likelihood that genetic epistasis should be detected. However, genetic epistasis requires not only interacting molecular pathways, but also allelic variation within interacting pathways of sufficient magnitude to provide a significant statistical interaction. Large QTL mapping studies find little evidence for epistatic interactions for specific developmental, architectural, and biochemical traits [56–58] although, as described previously, heterosis is greater for highly complex traits such as grain yield, traits for which quantitative genetic studies more often support the role of epistasis. In cases in which qualitative mutations have been introgressed into multiple genetic backgrounds, there is compelling evidence that expression is highly background-dependent. Therefore, it is logical by extension that genes of smaller effect should interact in the same way. However, the effect of individual genes/QTL must be of sufficient magnitude for interactions to be detectable within the constraints of specific experimental designs and population sizes. Understanding of the role of epistasis in heterosis and expression of other traits will continue to improve as molecular tools and statistical approaches advance. Current evidence suggests that there is much more to be learned about epistatic gene interactions underlying heterosis.

It is important to recognize that estimates of gene action are based on a logical framework of genes, alleles, and allelic effects (e.g., Falconer and Mackay [59]), and interpretations are only relevant in the context of that framework. In the next section, I will discuss molecular

mechanisms that are consistent with that framework. However, mechanisms of phenotypic variation due to locus-independent, genomewide mechanisms have been proposed and will be summarized later in this paper. Note that phenotypic variation due to these mechanisms will still be partitioned within the context of gene-specific models in variance component studies due to restrictions in the model, but may actually result from a more general mechanism.

6. Molecular Evidence Consistent with Quantitative Genetic Models

The concept of heterosis has evolved parallel to discoveries on the molecular basis of mutation, the control of transcription and translation, and the discovery of heritable chromatin-based allelic states. Quantitative genetic models underlying current breeding and variance partitioning models are based on heritable allelic variation that provides consistent effects within defined genetic and environmental contexts. An early and still prevalent model of alternative allelic states is the presence of every gene in all individuals of a species with an array of sequence variants that could confer minor to extreme functional consequences including intermediate-to-complete loss-of-function alleles. This concept is consistent with extensive single-nucleotide polymorphism (SNP), indel, and transposon variation found within and near genes when comparing individual genomes within species [60–65]. The discovery that plant genomes contain a large proportion of repetitive transposons raises the possibility for transposons to influence the expression of nearby genes including altering expression levels, producing ectopic gene expression, and producing allelic variation by introducing footprints following insertion and excision [66]. Recently, the growing realization of the importance of presence-absence variation (PAV) and copy number variation (CNV) supports the concept of pangenomes within species in which all individuals within a species may not contain a copy of all the genes found across the species [67–70]. Finally, heritable epialleles [71] provide a sequence-independent mechanism to produce altered expression levels that might be able to more rapidly revert to support rapid direct or natural evolutionary change.

All of these allele-generating mechanisms—SNPs, transposons, PAVs, and epialleles—are consistent with the hypothesis that locus-specific intra-allelic interactions with some degree of dominance are responsible for heterosis. For example, SNPs can reduce function by altering the activity or productivity of enzymes or by reducing the efficiency of transcription factor binding. Loss-of-function could result from SNPs producing nonsense alleles or altering splice junctions, or loss of transcript due to absence of a sequence or by epigenetic silencing. Alleles with reduced or complete loss of function can be accumulated in random-mating highly heterozygous populations of individuals. Upon inbreeding, homozygosity of deleterious alleles would result in loss of vigor (inbreeding depression) that would be restored by mating of genetically unrelated individuals.

Novel alleles occur in the context of chromosomal locations, and recent studies that define the nonlinearity of recombination event frequency across the chromosome [72] are consistent with observations of pseudooverdominance. Accumulation of mutations in centromeric regions with limited recombination results in quantitative genetic estimates of overdominance in variance analysis and QTL studies due to the high degree of persistent linkage disequilibrium in these regions. The potential of regions with limited recombination to harbor deleterious alleles that rarely have the opportunity to recombine is the basis of the concept of heterotic patterns used by plant breeders, is consistent with heterosis observed in genetically isolated natural and artificial populations, and provides a basis of the value of polyploidy to fix heterotic gene interactions by combining divergent but related genomes.

Heterotic patterns used by plant breeders [73] provide a useful conceptual model to discuss heterosis in isolated populations. Breeders have purposefully separated breeding lines into distinct groups (parental pools) and limited intermating between pools as a way to maximize the performance of hybrids between parents selected from the groups. Consider, for example, the possibility that a species with 10 chromosomes has a pair of loci on each of the 10 chromosomes 1 centimorgan apart in repulsion phase with dominant gene action. It would be relatively straightforward based on phenotype or genotype to develop two breeding pools that would be fixed for the complementing allelic pairs at each of those 10 positions producing full performance of the hybrids between the pools. However, gametes containing recombination events in each of the intervals would be required to produce an individual out of the founder population with favorable alleles at all 20 loci (10 pairs). In a single generation, this combination would occur at a frequency of $0.05^{10} = 1$ in 10 trillion individuals, more than 5 times the number of corn plants grown in the United States in any year. In reality, the situation is much more complex with multiple loci in repulsion phase in genomic regions of high and persistent LD making it logical to capture the performance potential of linkage blocks as opposed to trying to identify exceptionally rare recombinant types that resolve repulsion-phase linkages. This concept can be applied to geographically or genetically isolated populations. Inbreeding due to drift would lead to divergence of genomic blocks in high LD regions resulting in reduced overall performance. After many generations of separation, heterosis would be observed upon crossing the populations to each other due to complementation.

While the dominance hypothesis has been described by some as the "old view" of heterosis, it is consistent with the majority and diversity of results observed across species including predictable heritability for performance when populations are subjected to selection, estimates of gene action in controlled experiments, and recent information on the molecular basis of allelism. Nevertheless, it is possible that quantitative genetic models conceived in the early 1900s do not adequately capture all of the molecular mechanisms understood today, and there are at least anecdotal accounts of specific hybrids that perform beyond expectation based

on classical quantitative genetic models. These observations continue to spur research into molecular mechanisms, perhaps genomewide and locus-independent, that are needed to explain at least some component of heterosis.

7. Genomic Analysis of Heterosis

Phenotype is the result of the interpretation of genetic information through the processes of transcription, translation, and metabolism and development. Genomic studies have, therefore, assessed the transcriptome, proteome, metabolome, and related control mechanisms in inbreds and hybrids as an approach to evaluate the relationship between observed phenotypes and underlying molecular pathways. The simplest interpretation would be a direct relationship between molecular expression and observed phenotype, such that additive amounts of transcript would produce an intermediate phenotype. It is important to note that the connection between molecular measures and final phenotype will likely not be that clear, as in the tomato example of overdominance cited above [43] in which the intermediate transcriptional expression at the *SFT* locus resulted in overdominance for yield.

Transcriptome studies measure the relative total amount of transcript per locus, or can measure the relative contribution of each allele in hybrids. Both types of information are useful and complementary, but it is important to recognize that they are different measures of transcription and that neither provides information on transcript of an individual gene per cell. Genome-wide studies of the transcriptome in inbred versus hybrid parents reveal that a majority of genes are expressed in an additive manner [74–76], and a smaller proportion of genes show nonadditive expression of which a very small percent show expression outside the parental values (transcriptional overdominance or epistasis). Non-additive gene action could result from genetic and epigenetic intra-allelic interactions including paramutation, or from interallelic interactions (epistasis). One example of an epistatic interaction resulting in expression beyond parental values would be the complementation of alleles in a heterodimeric transcription factor that would result in transcriptional activation of a pathway in a hybrid that is not transcriptionally active in either parent due to absence of one component. It is notable that this type of epistatic interaction is rarely observed in genomewide transcriptome studies.

Overall transcription at a locus is a combined contribution from each parent. It is possible that an additive value of expression could result from a linear contribution of each parental allele in the hybrid relative to its expression in the inbred (*cis* control) or could be due to the heterozygote of a distant controlling factor modulating the level of expression (*trans* control). Stupar and Springer [74] evaluated the allelic contribution to expression in the hybrid across multiple loci and found that the majority of loci were controlled in *cis*. This is generally consistent with observations by Guo et al. [77] who studied genomewide allele-specific expression in maize hybrids and found primarily intermediate contributions from both parents with some loci exhibiting maternal

or paternal bias. In a related study, Guo et al. [78] reported that paternally biased expression was higher under the stress of high plant density and higher in an old hybrid versus a new hybrid indicating a potentially important environmental component to observed expression values.

Additive transcript levels of genes could result in nonadditive phenotypic performance in several ways. First, presence of a single favorable allele may be sufficient to provide protein function equivalent to the high-parent level even if both are expressed and the favorable allele is present in only one-half the amount. Second, additive expression levels could be observed in the hybrid in cases of a presence-absence allelic contrast in the parent with one parent having no expression and the other expressing a functional product. The hybrid may have only half the expression of the parent containing the gene, but that amount of expression could be sufficient to complement the deficiency due to the absence of the gene in the other parent. Therefore, the observed results are consistent with quantitative genetic observations based on phenotype. It is notable that the hybrid is generally a predictable combination of the inbred parents and that it does not exhibit genomewide luxuriant transcription levels which are not predictable by parental expression levels as suggested by some models [79].

Various studies have measured small RNA levels in inbreds and hybrids, some of which present a strong suggestion for the role of small RNAs in heterosis [80, 81]. A recent study in maize using Illumina sequencing and qPCR confirmation revealed that, as with gene transcription, small RNA levels are generally additive in the hybrid with amounts predictable based on the inbred [82]. It is possible that, as with genic transcription factors, additive interactions among different small RNAs could result in non-additive expression of the loci that they control, although this type of expression is a minority. An interesting finding in the Barber et al. [82] study was the observation that hybrid maize plants relative to their inbred parents, all containing the *mop1* mutation (a protein which is necessary for most 24nt small RNA production), were equally or more heterotic than nonmutant hybrids. This result indicates that this specific class of small RNAs is not required for heterotic phenotypic expression in maize hybrids.

Proteomic analysis is another approach that has been used to characterize molecular components of heterosis. Proteomic analysis of seedling roots of maize [83–86] and rice [87] indicates that non-additive expression of proteins in hybrids versus inbreds is more frequent than non-additive transcriptional variation. Dahal et al. [88] compared two heterotic maize hybrids to a non-heterotic hybrid. They found that proteins enriched in stress response and protein and carbon metabolism were differentially expressed in heterotic hybrids. Their results indicated that the degree of heterosis was correlated with the frequency of protein isoforms and/or modifications.

In summary, extensive genomic studies provide insights but no direct answers regarding the basis of heterosis. All modes of gene action—additivity, dominance, overdominance, and epistasis—are observed at the molecular level, but the interpretation of those molecular effects to

final phenotype remains complex and largely undefined. Overall, the results are consistent with the importance of specific allelic variants in the manifestation of heterosis and with the predictable inheritance of molecular phenotypes. However, some mechanisms have been proposed that are independent of allelic effects and rather are genomewide responses to genomic diversity. These potential mechanisms will be discussed in the following section.

8. Genomewide Models to Explain Heterosis

Heterotic expression of phenotypes is, in many instances, correlated with genetic distance [89–93]. While this is generally true, the relationship is clearest in the comparison of hybrids with similar adaptation, and which have been selected for productivity (summarized by Melchinger [94]). An example would be the collection of public and private, off-PVP maize inbreds released in the US that have been selected for performance in generally similar contexts. Within this group, there would be an expectation of a strong correlation between genetic diversity and performance based on the breeding method by which the lines were developed. As the genetic distance becomes greater, and complexities of adaptation are introduced, the relationship between performance and genetic diversity is lost. Therefore, genetic diversity *per se* is not the sole basis for heterosis. By extension, other mechanisms that generate diversity such as mutagenesis would not be expected to produce heterosis commensurate with the degree of divergence. Nevertheless, it has been postulated that the genome has mechanisms to sense diversity and the response to diversity can be translated into heterotic performance. Genomewide mechanisms are those considered to be gene/allele-independent. Note that, based on this definition, genomewide mechanisms would also be considered to be trait-independent producing heterosis for all traits to a similar degree. In general, heterosis across hybrids is not general but rather is trait-specific (summarized in Kaeppler [95]).

One genomewide mechanism that has been proposed as a basis of heterosis is changes in DNA methylation, or more broadly, chromatin state. Heritable epigenetic variation is a common attribute of plant genomes, likely more frequent than sequence variation (Becker and Weigel [96]). The possibility of directed, or at least more frequent, changes in DNA methylation in hybrids relative to their inbred progenitors is consistent with the potential stimulation of growth based on diversity *per se*. It is also commensurate with allele- and locus-specific observations of paramutation [97–99] in which the allelic interaction results in a heritable change in expression state, an observation inconsistent with the tenets of quantitative genetic theory. Recent studies of genomewide methylation analysis by sequencing inbreds and hybrids suggest that repeatable methylation changes upon hybridization, likely directed by small RNAs, may be somewhat common [100, 101], but more research is needed to understand the impact of these changes on gene expression and phenotype.

Sequence-based analysis of DNA methylation provides more detail than previous studies based on total proportion of 5-methyl cytosine in the genome, but studies based on proportion of methylated cytosines provide some intriguing hints about environmental influences on methylation changes, and potential differences among species. Tsfartis et al. [102] reported reduced levels of DNA methylation in hybrid relative to inbred maize plants and found the reductions to be related to stress (planting density). Furthermore, alterations in methylation were found to be heritable. Recently, Vergeer et al. [103] reported that inbreeding in *Scabiosa* is correlated with increased genomewide DNA methylation and methylation is reduced in hybrids. Furthermore, they report that application of a demethylating agent, 5-azacytidine, to inbreds restored productivity to the hybrid level. While 5-azacytidine has genomewide effects, it is not clear if the observed stimulation of vigor is a locus-specific effect, perhaps related with flowering. This result is in contrast to Shen et al. [101] who reported increases in DNA methylation in hybrids relative to inbred progenitors and reduced vigor in hybrids treated with a chemical that reduced methylation. Generally, in species exhibiting inbreeding depression, little evidence exists that neither DNA methylation or chromatin mutants, nor chemical treatments to reduce DNA methylation or alter histone modification, will stimulate vigor. In most cases, vigor would be expected to be reduced in these mutants and by these treatments.

Goff [104] proposed a model accounting from multigenic heterosis based on gains in energy efficiency due to protein processing in hybrids relative to inbreds. The model proposes that allelic choice available in hybrids but not inbreds provides hybrids the opportunity to detect and express preferentially the favorable allele. By minimizing expression of alleles that will require energy-intensive protein recycling, hybrids realize a synergistic growth benefit that begins to be realized during early growth with benefits accumulating throughout the life-cycle of the plant. This idea is consistent with the idea that diversity *per se* is not the basis of heterosis, but maximizing "quality" alleles in hybrids contributes to performance regardless of the function of those genes. It is in contrast to the observation that manifestation of heterosis is trait-dependent. Genomewide models of heterosis predict that vigor for all heterotic traits would benefit similarly.

9. Polyploidy, Aneuploidy, and Heterosis

Polyploidy provides a mechanism to capture heterotic gene combinations. In addition, the phenotypic consequences of gene copy number in polyploids and aneuploids, even those containing single alleles at all loci, may offer hints about mechanisms underlying heterosis [13, 20, 105].

Allopolyploids are formed by the union of distinct genomes in a single nucleus. The process of allopolyploidization can result from hybridization followed by somatic chromosome doubling or, more frequently, fertilization of unreduced gametes. Allelic complementation at common loci in the homoeologous genomes is fixed upon polyploidization, thereby fixing heterotic potential contributed

by the component species. This mechanism of capturing heterotic performance through the process of polyploidization is consistent with the dominance/overdominance/epistatic models described above. Furthermore, polyploids have additional opportunity for epistatic interactions due to potential segregation of interacting loci contributed by the component genomes as well as independent segregation of allelic variants at homoeologous loci.

An interesting observation in autopolyploids of progressive heterosis [106]. Progressive heterosis is the increase in performance of individuals as the probability of allelic diversity increases. Specifically, the level of performance is greater when more than two alleles at a locus are possible than when only two alleles can be present. The observation of progressive heterosis has alternatively been interpreted as consistent with pseudooverdominance due to repulsion phase linkage of dominant alleles [107] and as an argument against simple complementation and for higher-order intra-allelic interactions [12]. Across diploid species, the bulk of current evidence supports complementation (dominance) versus intra-allelic interactions (overdominance).

An intriguing phenotypic consequence of polyploidy and aneuploidy is the difference in performance due to the number of genomic complements, or to variation in doses of whole chromosomes or portions of chromosomes, and these consequences may have implications for heterosis [12, 108, 109]. These differences in performance can occur independent of any allelic diversity. Haploids in plants are generally lacking vigor, and doubled haploids (dihaploids) are as vigorous as sexually derived individuals of the same ploidy while being completely homozygous. In cases, where polyploid series have been produced, individuals of higher ploidy are often more vigorous than lower ploidy progenitors, although fertility is often compromised. Therefore, increased performance for traits such as forage yield is possible in the absence of allelic diversity simply by increasing DNA content per cell. On the other hand, altering the dosage of chromosomes or chromosome segments in aneuploids often reduces vigor and performance. In aneuploids, under- and over-representation of chromosome segments similarly results in reduced vigor. Therefore, pathways clearly exist across organisms to sense gene dosage [109], and the phenotypic consequences of polyploidy and aneuploidy are similar to differential performance of inbreds and hybrids. In the context of the dramatic presence/absence and copy number variation observed in many species, it is interesting to consider the possibility that dosage sensing is an allele-independent mechanism underlying heterosis. For example, consider that through segregation of PAV/CNV alleles, inbreds accumulate a specific level of average dosage imbalance across the genome due to segregation and result in reduction in vigor. Hybrids formed between crosses of inbred lines would have average gene copy number across the genome that would be less deviant than either inbred parent restoring vigor. From a breeding standpoint, if dosage imbalance is important in performance, selection based on performance would tend to minimize CNV in genomes, at least at loci subject to a dosage response.

10. Summary and Integration

It is clear that much remains to be learned about genome composition and the role of transcription, translation, and posttranslational mechanisms in interpreting genes into phenotype. While it is certain that future discoveries will explain more about the process of heterosis, it is my opinion that a new and undiscovered molecular mechanism is not needed to ultimately explain heterosis. Heterosis is greatest for highly complex traits composed of multiple component phenotypes. An accumulation of the effects of a large number of genes with small effects and some level of dominance, taken in the context of recombination across the genome, is sufficient to explain heterosis and is consistent with directed and natural evolution. Mechanistically, the undiscovered territory is the multiplicity of specific mechanisms by which the cumulative influence of a large number of allelic variants is manifested.

Discussions of heterosis are often confused by inconsistent separation of absolute measurement of performance (yield, productivity, etc.) versus true measures of heterosis which is the deviation of the performance of a hybrid individual or population from its parental progenitor. Performance of many traits has been shown to be inherited in an expected and repeatable manner, indicating that performance in the hybrid state cannot be the result of mechanisms that are not manifested through selection and inbreeding. Quantitative genetic models based on dominance and epistasis explain heterosis, observed phenotypic variation and are consistent with observations of reduced heterosis (deviation of hybrid performance from mean of inbreds) as the performance of hybrids is improved. Recent genomic studies which show that large regions of the genome have limited recombination, providing a mechanism for the accumulation of deleterious mutations that can only be resolved and purged in rare recombinant gametes. The increasing number of ways that deleterious alleles can be produced including SNPs, transposon insertions and signatures, PAV, and epiallelic variation provides new ways to account for formation of deleterious alleles. The bulk of available data are highly consistent with the dominance (complementation) hypothesis as the primary basis of heterosis. Furthermore, heterosis is of greatest magnitude in highly complex traits such as grain yield which is affected by many interacting developmental, metabolic, and environment response pathways supporting that a large number of genes, likely each with small effects, are cumulatively responsible in the context of interacting (epistatic) pathways to explain performance and heterosis. Diverse molecular mechanisms that interpret DNA sequence into phenotype will be involved, and research to characterize pathways and fundamental molecular mechanisms will be important to understand heterosis in the context of diverse phenotypes, each independently displaying heterosis in specific genetic contexts.

There is no missing, gene-independent, unifying mechanism to explain heterosis—heterosis is the result of the diversity of genes, pathways, and processes known and yet to be discovered. Specific examples may highlight one

mechanism or process in the context of a specific trait and genetic context, but those examples are just examples and do not overshadow the fact that extant natural variation is the resulting accumulation of the results of millennia of mutation and natural and artificial selection manifested in the organisms that we measure today. To say that there is no missing unifying mechanism is not meant to diminish the importance of fundamental research. Rather it is meant to highlight the importance of diverse fundamental experiments to ultimately understand biologically and economically important phenomena such as heterosis and to suggest that the final answer to the basis of heterosis will be the accumulation of results of many and diverse studies and not a singular, unifying, novel discovery.
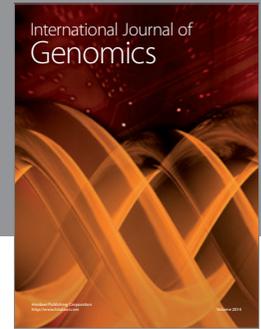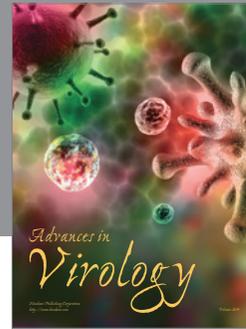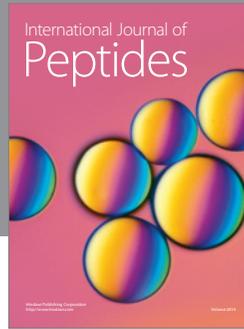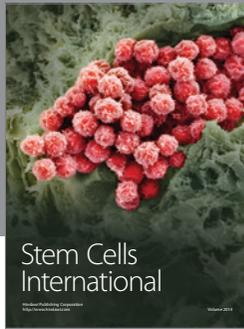
Acknowledgments

References

[1] M. B. Lewis, "Why are mixed-race people perceived as more attractive?" *Perception*, vol. 39, no. 1, pp. 136–138, 2010.

[2] M. A. Mingroni, "Resolving the IQ paradox: heterosis as a cause of the Flynn effect and other trends," *Psychological Review*, vol. 114, no. 3, pp. 806–829, 2007.

[3] M.A. Woodley, "Heterosis doesn't cause the Flynn effect: a critical examination of Mingroni (2007)," *Psychological Review*, vol. 118, no. 4, pp. 689–693, 2011.

[4] H. Campbell, A. D. Carothers, I. Rudan et al., "Effects of genome-wide heterozygosity on a range of biomedically relevant human quantitative traits," *Human Molecular Genetics*, vol. 16, no. 2, pp. 233–241, 2007.

[5] R. Khongsdier and N. Mukherjee, "Effects of heterosis on growth in height and its segments: a cross-sectional study of the Khasi girls in Northeast India," *Annals of Human Biology*, vol. 30, no. 5, pp. 605–621, 2003.

[6] S. Koziel, D. P. Danel, and M. Zareba, "Isolation by distance between spouses and its effect on children's growth in height," *American Journal of Physical Anthropology*, vol. 146, no. 1, pp. 14–19, 2011.

[7] V. Jakubec, "Productivity of crosses based on prolific breeds of sheep," *Livestock Production Science*, vol. 4, no. 4, pp. 379–392, 1977.

[8] B. Gjerde and T. Refstie, "Complete diallel cross between five strains of Atlantic salmon," *Livestock Production Science*, vol. 11, no. 2, pp. 207–226, 1984.

[9] O. Syrstad, "Heterosis in Bos taurus × Bos indicus crosses," *Livestock Production Science*, vol. 12, no. 4, pp. 299–307, 1985.

[10] G. Bittante, L. Gallo, and P. Montobbio, "Estimated breed additive effects and direct heterosis for growth and carcass traits of heavy pigs," *Livestock Production Science*, vol. 34, no. 1-2, pp. 101–114, 1993.

[11] J. S. Gavora, R. W. Fairfull, B. F. Benkel, W. J. Cantwell, and J. R. Chambers, "Prediction of heterosis from DNA fingerprints in chickens," *Genetics*, vol. 144, no. 2, pp. 777–784, 1996.

[12] J. A. Birchler, D. L. Auger, and N. C. Riddle, "In search of the molecular basis of heterosis," *Plant Cell*, vol. 15, no. 10, pp. 2236–2239, 2003.

[13] J. A. Birchler, H. Yao, and S. Chudalayandi, "Unraveling the genetic basis of hybrid vigor," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 35, pp. 12957–12958, 2006.

[14] J. A. Birchler, H. Yao, S. Chudalayandi, D. Vaiman, and R. A. Veitia, "Heterosis," *Plant Cell*, vol. 22, no. 7, pp. 2105–2112, 2010.

[15] J. F. Crow, "90 years ago: the beginning of hybrid maize.," *Genetics*, vol. 148, no. 3, pp. 923–928, 1998.

[16] D. N. Duvick, "Heterosis: feeding people and protecting resources," in *Genetics and Exploitation of Heterosis in Crop Plants*, J. G. Coors and S. Pandey, Eds., pp. 19–29, Crop Science Society of America, Madison, Wis, USA, 1999.

[17] Z. B. Lippman and D. Zamir, "Heterosis: revisiting the magic," *Trends in Genetics*, vol. 23, no. 2, pp. 60–66, 2007.

[18] C. W. Stuber, "Heterosis in plant breeding," *Plant Breeding Reviews*, vol. 12, pp. 227–251, 1994.

[19] N. M. Springer and R. M. Stupar, "Allelic variation and heterosis in maize: how do two halves make more than a whole?" *Genome Research*, vol. 17, no. 3, pp. 264–275, 2007.

[20] Z. J. Chen, "Molecular mechanisms of polyploidy and hybrid vigor," *Trends in Plant Science*, vol. 15, no. 2, pp. 57–71, 2010.

[21] H. K. Hayes, "Development of the heterosis concept," in *Heterosis*, J. W. Gowen, Ed., pp. 49–65, Iowa State College Press, Ames, Iowa, USA, 1952.

[22] G. F. Shull, "Beginnings of the heterosis concept," in *Heterosis*, J. W. Gowen, Ed., pp. 14–48, Iowa State College Press, Ames, Iowa, USA, 1952.

[23] I. L. Goldman, "Inbreeding and outbreeding in the development of a modern heterosis concept," in *Genetics and Exploitation of Heterosis in Crop Plants*, J. G. Coors and S. Pandey, Eds., pp. 7–18, Crop Science Society of America, Madison, Wis, USA, 1999.

[24] G. N. Collins, "The importance of broad breeding in corn," *USDA Bureau of Plant Industry Bulletin*, vol. 141, part 4, pp. 33–42, 1909.

[25] C. Darwin, *On the Various Contrivances by Which British and Foreign Orchids are Fertilised by Insects, and On the Good Effects of Intercrossing*, Murray, London, UK, 1862.

[26] E. M. East, "Inbreeding in corn," *Reports of the Connecticut Agricultural Experiment Station*, vol. 1907, pp. 419–428, 1908.

[27] E. M. East and D. F. Jones, *Inbreeding and Outbreeding*, J.B. Lippincott Co., Philadelphia, Pa, USA, 1919.

[28] D. F. Jones, "The effects of inbreeding and crossbreeding upon development," *Connecticut Agricultural Experiment Station Bulletin*, vol. 207, pp. 419–428, 1918.

[29] G. E. Morrow and F. D. Gardner, "Field experiments with corn, 1893," *University of Illinois Agricultural Experiment Station Bulletin*, vol. 31, pp. 333–360, 1894.

[30] G. F. Shull, "The composition of a field of maize," *American Rabbit Breeders Association*, vol. 5, pp. 51–59, 1908.

[31] D. Charlesworth and J. H. Willis, "The genetics of inbreeding depression," *Nature Reviews Genetics*, vol. 10, no. 11, pp. 783–796, 2009.

[32] S. A. Flint-Garcia, E. S. Buckler, P. Tiffin, E. Ersoz, and N. M. Springer, "Heterosis is prevalent for multiple traits in diverse maize germplasm.," *PloS one*, vol. 4, no. 10, p. e7433, 2009.

[33] A. R. Hallauer, M. J. Carena, and J. B. Miranda Filho, "Quantitative genetics in maize breeding," in *Chapter 10: Heterosis*, pp. 477–523, Springer, New York, NY, USA, 2010.

[34] R. H. Moll, M. F. Lindsey, and H. F. Robinson, "Estimates of genetic variances and the level of dominance in maize," *Genetics*, vol. 49, pp. 411–423, 1963.

[35] C. O. Gardner and J. H. Lonnquist, "Linkage and the degree of dominance of genes controlling quantitative characters in maize," *Agronomy Journal*, vol. 51, pp. 524–528, 1959.

[36] W. A. Russell, S. A. Eberhart, and U. A. Vega, "Recurrent selection for specific combining ability for yield in two maize populations," *Crop Science*, vol. 13, pp. 257–261, 1978.

[37] M. D. Edwards, C. W. Stuber, and J. F. Wendel, "Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action.," *Genetics*, vol. 116, no. 1, pp. 113–125, 1987.

[38] C. W. Stuber, S. E. Lincoln, D. W. Wolff, T. Helentjaris, and E. S. Lander, "Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers," *Genetics*, vol. 132, no. 3, pp. 823–839, 1992.

[39] G. I. Graham, D. W. Wolff, and C. W. Stuber, "Characterization of a yield quantitative trait locus on chromosome five of maize by fine mapping," *Crop Science*, vol. 37, no. 5, pp. 1601–1610, 1997.

[40] A. Lariepe, B. Mangin, S. Jasson et al., "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*, vol. 190, pp. 795–811, 2012.

[41] J. Xiao, J. Li, L. Yuan, and S. D. Tanksley, "Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers," *Genetics*, vol. 140, no. 2, pp. 745–754, 1995.

[42] D. Schwartz and W. J. Laughner, "A molecular basis for heterosis," *Science*, vol. 166, no. 3905, pp. 626–627, 1969.

[43] U. Krieger, Z. B. Lippman, and D. Zamir, "The flowering gene *SINGLE FLOWER TRUSS* drives heterosis for yield in tomato," *Nature Genetics*, vol. 42, no. 5, pp. 459–463, 2010.

[44] Y. Semel, J. Nissenbaum, N. Menda et al., "Overdominant quantitative trait loci for yield and fitness in tomato," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 35, pp. 12981–12986, 2006.

[45] D. P. Wolf and A. R. Hallauer, "Triple testcross analysis to detect epistasis in maize," *Crop Science*, vol. 37, no. 3, pp. 763–770, 1997.

[46] B. Kusterer, J. Muminovic, H. F. Utz et al., "Analysis of a triple testcross design with recombinant inbred lines reveals a significant role of epistasis in heterosis for biomass-related traits in arabidopsis," *Genetics*, vol. 175, no. 4, pp. 2009–2017, 2007.

[47] A. E. Melchinger, H. P. Piepho, H. F. Utz et al., "Genetic basis of heterosis for growth-related traits in Arabidopsis investigated by testcross progenies of near-isogenic lines reveals a significant role of epistasis," *Genetics*, vol. 177, no. 3, pp. 1827–1837, 2007.

[48] J. C. Reif, B. Kusterer, H. P. Piepho et al., "Unraveling epistasis with triple testcross progenies of near-isogenic lines," *Genetics*, vol. 181, no. 1, pp. 247–257, 2009.

[49] S. B. Yu, J. X. Li, C. G. Xu et al., "Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 17, pp. 9226–9231, 1997.

[50] Z.-K. Li, L. J. Luo, H. W. Mei et al., "Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components," *Genetics*, vol. 158, no. 4, pp. 1755–1771, 2001.

[51] J. Hua, Y. Xing, W. Wu et al., "Single-locus heterotic effects and dominance by dominance interactions can adequately explain the genetic basis of heterosis in an elite rice hybrid," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 5, pp. 2574–2579, 2003.

[52] S. B. Rood, P. B. Kaufman, H. Abe, and R. P. Pharis, "Gibberellins and gravitropism in maize shoots: endogenous gibberellin-like substances and movement and metabolism of [3H]Gibberellin A20," *Plant physiology*, vol. 83, pp. 645–651, 1987.

[53] S. B. Rood, R. P. Pharis, M. Koshioka, and D. J. Major, "Gibberellins and heterosis in maize: I. Endogenous gibberellin-like substances," *Plant Physiology*, vol. 71, pp. 693–644, 1983.

[54] S. B. Rood, R. I. Buzzell, L. N. Mander, D. Pearce, and R. P. Pharis, "Gibberellins: a phytohormonal basis for heterosis in Maize," *Science*, vol. 241, no. 4870, pp. 1216–1218, 1988.

[55] Q. Ma, P. Hedden, and Q. Zhang, "Heterosis in rice seedlings: its relationship to gibberellin content and expression of gibberellin metabolism and signaling genes," *Plant Physiology*, vol. 156, no. 4, pp. 1905–1920, 2011.

[56] E. S. Buckler, J. B. Holland, P. J. Bradbury et al., "The genetic architecture of maize flowering time," *Science*, vol. 325, no. 5941, pp. 714–718, 2009.

[57] J. P. Cook, M. D. McMullen, J. B. Holland et al., "Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels," *Plant Physiology*, vol. 158, no. 2, pp. 824–834, 2012.

[58] F. Tian, P. J. Bradbury, P. J. Brown et al., "Genome-wide association study of leaf architecture in the maize nested association mapping population," *Nature Genetics*, vol. 43, no. 2, pp. 159–162, 2011.

[59] D. S. Falconer and T. F. C. Mackay, *Introduction to Quantitative Genetics*, Pearson, Prentice Hall, Edinburgh, Scotland, 4th edition, 1996.

[60] J.-M. Chia, C. Song, P. J. Bradbury et al., "Maize HapMap V2-Capturing variation in a genome in flux," *Nature Genetics*, vol. 44, pp. 803–807, 2012.

[61] M. B. Hufford, X. Xu, J. vanHeerwaarden et al., "Population genomics of domestication and improvement in maize," *Nature Genetics*, vol. 44, pp. 808–811, 2012.

[62] Y. Jiao, H. Zhao, L. Ren et al., "Genome-wide genetic changes during modern breeding of maize," *Nature Genetics*, vol. 44, pp. 812–815, 2012.

[63] J. Lai, R. Li, X. Xu et al., "Genome-wide patterns of genetic variation among elite maize inbred lines," *Nature Genetics*, vol. 42, no. 11, pp. 1027–1030, 2010.

[64] X. Xu, X. Liu, S. Ge et al., "Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes," *Nature Biotechnology*, vol. 30, pp. 105–111, 2011.

[65] M. Nordborg, T. T. Hu, Y. Ishino et al., "The pattern of polymorphism in *Arabidopsis thaliana*," *PLoS Biology*, vol. 3, no. 7, p. e196, 2005.

[66] M. G. Kidwell and D. Lisch, "Transposable elements as sources of variation in animals and plants," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 15, pp. 7704–7711, 1997.

[67] H. Fu and H. K. Dooner, "Intraspecific violation of genetic colinearity and its implications in maize," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 14, pp. 9573–9578, 2002.

[68] S. Brunner, K. Fengler, M. Morgante, S. Tingey, and A. Rafalski, "Evolution of DNA sequence nonhomologies among maize inbreds," *Plant Cell*, vol. 17, no. 2, pp. 343–360, 2005.

[69] N. M. Springer, K. Ying, Y. Fu et al., "Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content," *PLoS Genetics*, vol. 5, no. 11, Article ID e1000734, 2009.

[70] C. N. Hansey, B. Vaillancourt, R. S. Sekhon, N. de Leon, S. M. Kaeppler, and C. R. Buell, "Maize (*Zea mays* L.) genome diversity revealed by RNA-sequencing," *PLoS One*, vol. 7, no. 3, Article ID e33017, 2012.

[71] I. Makarevitch, R. M. Stupar, A. L. Iniguez et al., "Natural variation for alleles under epigenetic control by the maize chromomethylase Zmet2," *Genetics*, vol. 177, no. 2, pp. 749–760, 2007.

[72] M. A. Gore, J. M. Chia, R. J. Elshire et al., "A first-generation haplotype map of maize," *Science*, vol. 326, no. 5956, pp. 1115–1117, 2009.

[73] W. F. Tracy and M. A. Chandler, "The historical and biological basis of the concept of heterotic patterns in corn belt dent maize," in *Hallauer Plant Breeding Symposium*, K. R. Lamkey and M. Lee, Eds., pp. 219–233, Blackwell, Ames, Iowa, USA, 2006.

[74] R. M. Stupar and N. M. Springer, "Cis-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid," *Genetics*, vol. 173, no. 4, pp. 2199–2210, 2006.

[75] R. A. Swanson-Wagner, Y. Jia, R. DeCook, L. A. Borsuk, D. Nettleton, and P. S. Schnable, "All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 18, pp. 6805–6810, 2006.

[76] R. A. Swanson-Wagner, R. Decook, Y. Jia et al., "Paternal dominance of trans-eQTL influences gene expression patterns in maize hybrids," *Science*, vol. 326, no. 5956, pp. 1118–1120, 2009.

[77] M. Guo, M. A. Rupe, C. Zinselmeier, J. Habben, B. A. Bowen, and O. S. Smith, "Allelic variation of gene expression in maize hybrids," *Plant Cell*, vol. 16, no. 7, pp. 1707–1716, 2004.

[78] M. Guo, M. A. Rupe, X. Yang et al., "Genome-wide transcript analysis of maize hybrids: allelic additive gene expression and yield heterosis," *Theoretical and Applied Genetics*, vol. 113, no. 5, pp. 831–845, 2006.

[79] M. Freeling, M. R. Woodhouse, S. Subramaniam, G. Turco, D. Lisch, and J. C. Schnable, "Fractionation mutagenesis and similar consequences of mechanisms removing dispensable or less-expressed DNA in plants," *Current Opinion in Plant Biology*, vol. 15, no. 2, pp. 131–139, 2012.

[80] D. Ding, Y. Wang, M. Han et al., "MicroRNA transcriptomic analysis of heterosis during maize seed germination," *PLoS ONE*, vol. 7, no. 6, Article ID e39578, 2012.

[81] D. W. Ng, J. Lu, and Z. J. Chen, "Big roles for small RNAs in polyploidy, hybrid vigor, and hybrid incompatibility," *Current Opinion in Plant Biology*, vol. 15, no. 2, pp. 154–161, 2012.

[82] W. T. Barber, W. Zhang, H. Win et al., "Repeat associated small RNAs vary among parents and following hybridization in maize," *Proceedings of the National Academy of Sciences*, vol. 109, no. 26, pp. 10444–10449, 2012.

[83] A. Paschold, C. Marcon, N. Hoecker, and F. Hochholdinger, "Molecular dissection of heterosis manifestation during early maize root development," *Theoretical and Applied Genetics*, vol. 120, no. 2, pp. 383–388, 2010.

[84] N. Hoecker, B. Keller, N. Muthreich et al., "Comparison of maize (*Zea mays* L.) F1-hybrid and parental inbred line primary root transcriptomes suggests organ-specific patterns of nonadditive gene expression and conserved expression trends," *Genetics*, vol. 179, no. 3, pp. 1275–1283, 2008.

[85] N. Hoecker, T. Lamkemeyer, B. Sarholz et al., "Analysis of nonadditive protein accumulation in young primary roots of a maize (*Zea mays* L.) F1-hybrid compared to its parental inbred lines," *Proteomics*, vol. 8, no. 18, pp. 3882–3894, 2008.

[86] C. Marcon, A. Schützenmeister, W. Schütz, J. Madlung, H. P. Piepho, and F. Hochholdinger, "Nonadditive protein accumulation patterns in maize (*Zea mays* L.) hybrids during embryo development," *Journal of Proteome Research*, vol. 9, no. 12, pp. 6511–6522, 2010.

[87] W. Wang, B. Meng, X. Ge et al., "Proteomic profiling of rice embryos from a hybrid rice cultivar and its parental lines," *Proteomics*, vol. 8, no. 22, pp. 4808–4821, 2008.

[88] D. Dahal, B. P. Mooney, and K. J. Newton, "Specific changes in total and mitochondrial proteomes are associated with higher levels of heterosis in maize hybrids," *Plant Journal*, vol. 72, no. 1, pp. 70–83, 2012.

[89] R. H. Moll, W. S. Salhuana, and H. F. Robinson, "Heterosis and genetic diversity in variety crosses of maize," *Crop Science*, vol. 2, no. 3, pp. 197–198, 1962.

[90] A. E. Melchinger, M. M. Messmer, M. Lee, W. L. Woodman, and K. R. Lamkey, "Diversity and relationships among U.S. inbreds revealed by restriction fragment length polymorphisms," *Crop Science*, vol. 31, no. 3, pp. 669–678, 1991.

[91] R. Bernardo, "Relationship between single-cross performance and molecular marker heterozygosity," *Theoretical and Applied Genetics*, vol. 83, no. 5, pp. 628–634, 1995.

[92] R. Bernardo, "Prediction of maize single-cross performance using RFLPs and information from related hybrids," *Crop Science*, vol. 34, no. 1, pp. 20–25, 1994.

[93] P. A. Marsan, P. Castiglioni, F. Fusari, M. Kuiper, and M. Motto, "Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers," *Theoretical and Applied Genetics*, vol. 96, no. 2, pp. 219–227, 1998.

[94] A. E. Melchinger, "Genetic diversity and heterosis," in *Genetics and Exploitation of Heterosis in Crop Plants*, J. G. Coors and S. Pandey, Eds., pp. 99–118, Crop Science Society of America, Madison, Wis, USA, 1999.

[95] S. Kaeppler, "Heterosis: one boat at a time, or a rising tide?" *New Phytologist*, vol. 189, no. 4, pp. 900–902, 2011.

[96] C. Becker and D. Weigel, "Epigenetic variation: origin and transgenerational inheritance," *Current Opinion in Plant Biology*, vol. 15, no. 1, pp. 1–6, 2012.

[97] R. A. Brink, "A genetic change associated with the R locus in maize which is directed and potentially reversible," *Genetics*, vol. 41, no. 4, pp. 872–889, 1956.

[98] V. L. Chandler and M. Stam, "Chromatin conversations: mechanisms and implications of paramutation," *Nature Reviews Genetics*, vol. 5, no. 7, pp. 532–544, 2004.

[99] J. B. Hollick and N. M. Springer, "Epigenetic phenomena and epigenomics in maize," in *Epigenomics*, A. C. Ferguson-Smith, J. M. Greally, and R. A. Martienssen, Eds., pp. 119–147, Springer Media B.V., Houton, The Netherlands, 2009.

[100] R. K. Chodavarapu, S. Feng, B. Ding et al., "Transcriptome and methylome interactions in rice hybrids," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 30, pp. 12040–12045, 2012.

[101] H. Shen, H. He, J. Li et al., "Genome-wide analysis of DNA methylation and gene expression changes in two *Arabidopsis ecotypes* and their reciprocal hybrids," *Plant Cell*, vol. 24, no. 3, pp. 875–892, 2012.

[102] A. S. Tsfartis, M. Kafka, A. Polidoros, and E. Tani, "Epigentic changes in maize DNA and heterosis," in *Genetics and Exploitation of Heterosis in Crop Plants*, J. G. Coors and S. Pandey, Eds., pp. 195–203, Crop Science Society of America, Madison, Wis, USA, 1999.

[103] P. Vergeer, N. Wagemaker, and N. J. Ouborg, "Evidence for an epigenetic role in inbreeding depression," *Biology Letters*, vol. 8, no. 5, pp. 798–801, 2012.

[104] S. A. Goff, "A unifying theory for general multigenic heterosis: energy efficiency, protein metabolism, and implications for molecular breeding," *New Phytologist*, vol. 189, no. 4, pp. 923–937, 2011.

[105] Z. J. Chen, "Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids," *Annual Review of Plant Biology*, vol. 58, pp. 377–406, 2007.

[106] R. W. Groose, L. E. Talbert, W. P. Kojis, and E. T. Bingham, "Progressive heterosis in autotetraploid alfalfa: studies using two types of inbreds," *Crop Science*, vol. 29, no. 5, pp. 1173–1177, 1989.

[107] E. T. Bingham, R. W. Groose, D. R. Woodfield, and K. K. Kidwell, "Complementary gene interactions in alfalfa are greater in autotetraploids than diploids," *Crop Science*, vol. 34, no. 4, pp. 823–829, 1994.

[108] J. A. Birchler, N. C. Riddle, D. L. Auger, and R. A. Veitia, "Dosage balance in gene regulation: biological implications," *Trends in Genetics*, vol. 21, no. 4, pp. 219–226, 2005.

[109] J. A. Birchler and R. A. Veitia, "Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 37, pp. 14746–14753, 2012.