



Heterosis Is a Systemic Property Emerging From Non-linear Genotype-Phenotype Relationships: Evidence From in Vitro Genetics and Computer Simulations

Julie Fievet, Thibault Nidelet, Christine Dillmann, Dominique de Vienne

► To cite this version:

Julie Fievet, Thibault Nidelet, Christine Dillmann, Dominique de Vienne. Heterosis Is a Systemic Property Emerging From Non-linear Genotype-Phenotype Relationships: Evidence From in Vitro Genetics and Computer Simulations. *Frontiers in Genetics*, 2018, 9, 26 p. 10.3389/fgene.2018.00159 . hal-01837851

HAL Id: hal-01837851

<https://hal.science/hal-01837851>

Submitted on 26 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Heterosis Is a Systemic Property Emerging From Non-linear Genotype-Phenotype Relationships: Evidence From *in Vitro* Genetics and Computer Simulations

Julie B. Fiévet^{1†}, Thibault Nidelet^{2†}, Christine Dillmann¹ and Dominique de Vienne^{1*}

¹ GQE–Le Moulon, INRA, Centre National de la Recherche Scientifique, AgroParisTech, Université Paris-Sud, Gif-sur-Yvette, France, ² Sciences Pour l’Enologie, INRA, Université de Montpellier, Montpellier, France

OPEN ACCESS

Edited by:

Dana C. Crawford,
Case Western Reserve University,
United States

Reviewed by:

Bethany Wolf,
Medical University of South Carolina,
United States
Diddahally Govindaraju,
Albert Einstein College of Medicine,
United States

*Correspondence:

Dominique de Vienne
dominique.de-vienne@inra.fr

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Applied Genetic Epidemiology,
a section of the journal
Frontiers in Genetics

Received: 13 November 2017

Accepted: 17 April 2018

Published: 15 May 2018

Citation:

Fiévet JB, Nidelet T, Dillmann C and
de Vienne D (2018) Heterosis Is a
Systemic Property Emerging From
Non-linear Genotype-Phenotype
Relationships: Evidence From *in Vitro*
Genetics and Computer Simulations.
Front. Genet. 9:159.
doi: 10.3389/fgene.2018.00159

Heterosis, the superiority of hybrids over their parents for quantitative traits, represents a crucial issue in plant and animal breeding as well as evolutionary biology. Heterosis has given rise to countless genetic, genomic and molecular studies, but has rarely been investigated from the point of view of systems biology. We hypothesized that heterosis is an emergent property of living systems resulting from frequent concave relationships between genotypic variables and phenotypes, or between different phenotypic levels. We chose the enzyme-flux relationship as a model of the concave genotype-phenotype (GP) relationship, and showed that heterosis can be easily created in the laboratory. First, we reconstituted *in vitro* the upper part of glycolysis. We simulated genetic variability of enzyme activity by varying enzyme concentrations in test tubes. Mixing the content of “parental” tubes resulted in “hybrids,” whose fluxes were compared to the parental fluxes. Frequent heterotic fluxes were observed, under conditions that were determined analytically and confirmed by computer simulation. Second, to test this model in a more realistic situation, we modeled the glycolysis/fermentation network in yeast by considering one input flux, glucose, and two output fluxes, glycerol and acetaldehyde. We simulated genetic variability by randomly drawing parental enzyme concentrations under various conditions, and computed the parental and hybrid fluxes using a system of differential equations. Again we found that a majority of hybrids exhibited positive heterosis for metabolic fluxes. Cases of negative heterosis were due to local convexity between certain enzyme concentrations and fluxes. In both approaches, heterosis was maximized when the parents were phenotypically close *and* when the distributions of parental enzyme concentrations were contrasted and constrained. These conclusions are not restricted to metabolic systems: they only depend on the concavity of the GP relationship, which is commonly observed at various levels of the phenotypic hierarchy, and could account for the pervasiveness of heterosis.

Keywords: heterosis, genotype-phenotype map, metabolic network, enzyme variability, non-linear processes, mathematical modeling

INTRODUCTION

The relationship between genetic polymorphism and phenotypic variation is a major issue in many branches of biology. It is usually referred to as the genotype-phenotype (GP) relationship, GP correlation or GP map (Lewontin, 1974). Strictly speaking, the genotype corresponds to the DNA level, whereas the phenotype can correspond to any trait observed or measured at any organizational level from transcription rate to fitness. In fact, following S. Wright (Wright, 1934), the GP relationship can also refer to the relationship between genetically variable phenotypic traits (e.g., enzyme activities) and more integrated phenotypic traits, i.e., traits measured at higher levels of the biological hierarchy (e.g., fluxes through networks).

Almost inevitably, the relationship between parameters at different organizational levels is non-linear, because the cellular processes that shape biological functions and structures are highly intertwined. The kinetics of biochemical and molecular reactions, transport and signaling mechanisms, positive and negative feedbacks, growth kinetics, etc., are intrinsically non-linear. Thus, from allosteric regulation to fitness landscapes, from genetic regulatory circuits to developmental processes, there are myriads of examples showing complex relationships between genotype and phenotypes, or between phenotypic levels (Kauffman, 1993; Heinrich and Schuster, 1996). These GP relationships display a large diversity of shapes (Rendel, 1962; Pryciak, 2008; Felix and Barkoulas, 2015), but two prevail to a large extent:

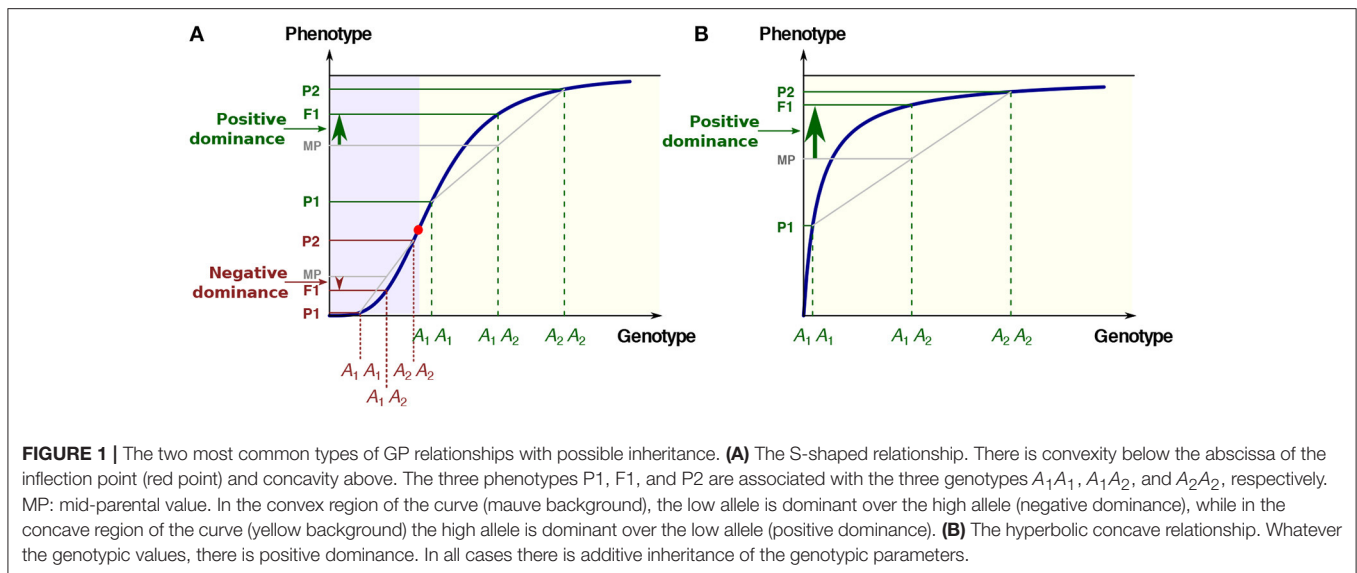
- (i) The sigmoidal relationship (**Figure 1A**). Cooperativity and synergy underlie such behavior (Carey, 1998; Kazemian et al., 2013), which is classically modeled using Hill functions. For instance in *Drosophila*, transcription of the *hunchback* gene in response to the gradient of Bicoid transcription factor concentration follows a typical S-shaped curve (Perry et al., 2012). Examples at other phenotypic levels exist, such as the relationship between the expression level of various genes and growth rate, a proxy for fitness, in *Escherichia coli* (Perfeito et al., 2011; Rodrigues et al., 2016) and in *Saccharomyces cerevisiae* (Rest et al., 2013).
- (ii) The concave relationship (**Figure 1B**), the archetype of which is the enzyme-flux relationship. The long-standing Metabolic Control Theory (MCT; reviewed in Fell, 2007) is probably the most elaborated corpus to describe this particular, but essential, case of genotype-phenotype relationship. Both theoretical developments and experimental observations show that increasing from zero the value of an enzyme parameter usually results first in an increase of the flux value, then in attenuation of the response and finally in saturation. The phenotypic response resembles a rectangular hyperbola, and at the asymptotic upper limit there is robustness of the phenotype toward genotypic variation.

Such concave relationships are common at other organizational levels. For instance, it is observed during transcriptional regulation when transcription factors do not bind to DNA cooperatively (Giorgetti et al., 2010). In yeast, the rate of

protein synthesis in response to variation in abundance of some 20 different translation factors is concave in most cases (Firczuk et al., 2013), and can be modeled using a deterministic model (Poker et al., 2014; Zarai et al., 2014). Other examples show the effects of variable gene/protein expression on the fitness of an organism. Izard et al. (2015), by engineering the control of RNA polymerase expression in *E. coli*, consistently observed a saturation curve for growth rate in diverse media. In mitochondrial diseases, the relationship between the proportion of mtDNA carrying pathogenic mutations and various cellular responses—mitochondrial protein synthesis rate, respiratory chain complex activity, mitochondrial respiration or ATP synthesis and growth rate—is concave, with a large plateau. Thus, an effect will only be detectable above a threshold value of the proportion of deleterious mtDNA (reviewed in Rossignol et al., 2003). In Parkinson's disease, the symptoms do not begin to appear until about 80% of neurons in the substantia nigra, which produce dopamine, have died. This is consistent with the hyperbola-shaped relationship between the proportion of alive neurons and the concentration of extra-cellular dopamine (Nijhout et al., 2015). In a wide range of species, allometry also provides relevant examples of non-linear relationships, such as between organismal mass and whole-organism metabolic rate, but also life span, tree trunk diameter, blood circulation time, etc. (West et al., 1999). Allometric equations are power laws whose scaling exponents are lower than 1, resulting in concave relationships.

In the above-mentioned studies, only one genotypic (or phenotypic) parameter is assumed to be variable. In fact, most integrated phenotypic traits are polygenic, i.e., they are potentially affected by a very large number of parameters, giving rise to complex and rugged phenotypic landscapes. However, from the previous considerations, one can expect global concavity of these multi-dimensional surface responses. Indeed, within the MCT framework, Dykhuizen et al. (1987) showed that the fitness response of *E. coli* to variation of both β -galactoside permease and β -galactosidase activity is a two-dimensional hyperbolic-like surface. Similarly, Nijhout et al. (2003) analyzed quantitatively the pairwise effects of various components of the MAPK signaling cascade on the MAPK-PP output and observed largely concave response surfaces with saturation. MacLean (2010) also found a concave relationship between growth rate and the joint variation of transcription and translation mediated by streptomycin and rifampicin, respectively, in *Pseudomonas aeruginosa*. Another interesting example is the allometric relationship described in a population of recombinant inbred lines of *Arabidopsis thaliana* between, on the one hand, photosynthetic rate and, on the other hand, dry mass per area and age at flowering (Vasseur et al., 2012).

Whether the phenotypic response is sigmoidal or concave, there is often a horizontal asymptote which accounts for the strong asymmetry of the phenotypic response, depending on whether the value of the genotypic variable is increased or decreased. Thus, in metabolic engineering it is usually much easier to decrease than to increase flux and metabolite concentrations (Moreno-Sanchez et al., 2008). At the genome-wide level, this asymmetry was recognized early on through the



classical work of Lindsley et al. (1972) on the effects of segmental aneuploidy in *Drosophila*. Using *Y*-autosome translocations with defined breakpoints, the authors produced a large number of small duplications and deficiencies spanning chromosomes 2 and 3. Examining the phenotypic effects of having the same chromosomal region present in one, two or three copies, the authors found that monosomy had a much greater effect on viability than trisomy, revealing a strong asymmetry of the response to gene dosage. “The pervasive robustness in biological systems” (Felix and Barkoulas, 2015), which has been observed at all phenotypic levels (transcript, protein and metabolic abundance, metabolic flux, chemotaxis, cell cycle period, cell fate patterning, cell survival, etc.), most likely stems from this saturation effect. In some cases the mechanisms of robustness have been unraveled (Barkai and Leibler, 1997; Eldar et al., 2004; Wagner, 2013), and a widespread link is thought to exist between robustness and network complexity (Martin and Wagner, 2008; Whitacre, 2012), even though robustness could be partly adaptive (Ho and Zhang, 2016).

In addition to robustness, the “diminishing return” curves could account for various fundamental and apparently unrelated observations. The best-known example is the dominance of “high” over “low” alleles for metabolic genes, which was first pointed out by Wright (1934), then formalized in detail by Kacser and Burns (1981) in a seminal paper. Due to concavity, the phenotypic value of the heterozygote exceeds the homozygote mean value, and alleles with large deleterious homozygous effects are more recessive than alleles of weaker effect (Figure 1). The metabolic concave relationship also has major implications regarding the evolution of selective neutrality (Hartl et al., 1985), the variability of enzyme activity in populations under mutation-selection balance (Clark, 1991), the relationship between flux and fitness (Dykhuizen et al., 1987), the epistasis between deleterious mutations (Szathmari, 1993), the response to directional selection (Keightley, 1996), the distribution of QTL effects (Bost et al., 1999, 2001), the dynamics of retention/loss of metabolic

genes after whole genome duplications (Gout et al., 2009) and the evolution of genes in branched pathways (Rausher, 2013).

The model of positive dominance of Wright, Kacser and Burns is not restricted to metabolic genes, since it is valid whenever there is concavity of the GP relationship. In the case of a sigmoidal response, dominance of deleterious mutations may occur, depending on the allelic values relative to the inflection point of the curve (Gilchrist and Nijhout, 2001; Bost and Veitia, 2014) (Figure 1A). It would be quite a laborious task to try to estimate the respective proportions of positive and negative dominance in the literature. Nevertheless, the fact remains that there is a marked bias toward the dominance of “high” alleles. As early as 1928, Fisher (1928) noted that, in *Drosophila*, 208 out of 221 autosomal or sex-linked deleterious mutations were recessive. Similar proportions have been observed in all species studied to date, including in artificial diploids of the normally haploid *Chlamydomonas* (Orr, 1991). More recently, genome-wide estimates from collections of deletion strains in yeast have been published (Steinmetz et al., 2002a; Deutschbauer et al., 2005). Whereas, 891 genes (20% of the genome) contributed to slow growth as homozygotes, only 184 (3% of the genome) were haploinsufficient (Deutschbauer et al., 2005). The mean values of the dominance coefficients were found to be $h = 0.25$ or less, which means that null alleles were on average recessive (Agrawal and Whitlock, 2011; Manna et al., 2011). Finally, predominant recessivity of deleterious mutations is fully consistent with common observations in population genetics, such as inbreeding depression (Charlesworth and Willis, 2009).

In this paper, we show that generalizing this dominance model to the multilocus case directly leads to a simple model of heterosis, i.e., a model that accounts for the common superiority of hybrids over their parents for quantitative traits (Shull, 1914; Jones, 1918; Gowen, 1952). Described more than 250 years ago (cited in Roberts, 1965) and studied in detail by Darwin (1876), hybrid vigor is of great importance for food production because it affects complex traits such as growth rate, biomass, fertility,

disease resistance, drought tolerance, etc. In crops, it is currently used to derive hybrid varieties with exceptional performances. For instance in maize, the hybrid grain yield can exceed by more than 100% the mean parental yield (Sprague, 1983; Duvick, 2001). In rice, heterosis is about 20% for grain yield and about 10% for plant height (Xiao et al., 1995). In livestock and farms animals, such as cattle, swine, sheep and poultry, heterosis is also present and commonly exploited (Dickerson, 1973; Yang et al., 1999). This phenomenon is not restricted to species of agronomic interest. In the model plant *A. thaliana*, there is heterosis for biomass, rosette diameter and flowering date (Barth et al., 2003; Rohde et al., 2004; Ni et al., 2009). In animals, it has been described for diverse traits in drosophila (Houle, 1989), frog (Hotz et al., 1999), Pacific oyster (Launey and Hedgecock, 2001), owned dogs (O'Neill et al., 2013), etc. Heterosis has also been found in microorganisms, for instance for growth rate in *Neurospora* (Emerson et al., 1952) and *S. cerevisiae* (Steinmetz et al., 2002b). In the latter case it may be spectacular, with growth rate one order of magnitude higher in the hybrid than in the best parent.

Theoretical models of heterosis have been proposed based on the functioning of gene circuits (Omholt et al., 2000) or metabolic systems (de Vienne et al., 2001; Fiévet et al., 2010). In the latter case the authors unraveled the links between heterosis, dominance and epistasis within the MCT framework. In support of this model, we show here that heterosis can easily be created in the laboratory. We relied on two complementary approaches: (i) Test-tube genetic experiments. We crossed *in vitro* “parents” that differed in the concentration of four successive glycolysis enzymes, and obtained “hybrids” displaying heterosis for flux through the pathway; (ii) *In silico* genetics. We crossed parents that differed in the concentration of 11 enzymes from the carbon metabolism network in *S. cerevisiae*, and we computed the input flux and two output fluxes in parents and hybrids using a system of differential equations. Heterotic fluxes were frequently observed. The conditions for having heterosis were analyzed using computer simulations and mathematically using the theory of concave functions. Small phenotypic difference between parents, together with contrasted and constrained parental enzyme concentrations, proved to be reliable predictors of heterosis. These conclusions are qualitatively valid beyond metabolic networks because they only depend on the shape of the GP relationship. Thus, the so-called “mysterious” phenomenon of heterosis appears to be a systemic property emerging from the non-linearity that exists at various levels of the biological hierarchy.

MATERIALS AND METHODS

Theoretical Background

Heterosis Indices

Let the relationship between a phenotypic variable z and a series of n genotypic variables x_j :

$$z = f(x_1, \dots, x_n).$$

Consider two individuals P1 and P2 with distributions of genotypic values $\{x_{j1}\}_{j=1,\dots,n}$ and $\{x_{j2}\}_{j=1,\dots,n}$, respectively, and their corresponding phenotypic values:

$$z_1 = f(x_{11}, \dots, x_{n1}) \text{ and } z_2 = f(x_{12}, \dots, x_{n2})$$

The hybrid P1*P2 has the genotypic distribution $\{x_{j,1*2}\}_{j=1,\dots,n}$ and the phenotypic value:

$$z_{1*2} = f(x_{1,1*2}, \dots, x_{n,1*2}).$$

When comparing z_{1*2} with the mean of z_1 and z_2 , heterosis can be observed. In its most common definition, heterosis corresponds to a positive departure of the hybrid value from the mean parental value for quantitative traits. However, negative heterosis (hybrid inferiority) also exists, and thus five situations of inheritance can be defined (Stupar et al., 2008; Zorgo et al., 2012; Blein-Nicolas et al., 2015) (Supplementary material, Figure S1). Following our notations, we have:

- Best-Parent Heterosis (BPH): $z_{1*2} > \max(z_1, z_2)$
- positive Mid-Parent Heterosis (+MPH): $(z_1 + z_2)/2 < z_{1*2} \leq \max(z_1, z_2)$
- Additivity (ADD): $z_{1*2} = (z_1 + z_2)/2$
- negative Mid-Parent Heterosis (–MPH): $\min(z_1, z_2) \leq z_{1*2} < (z_1 + z_2)/2$
- Worst-Parent Heterosis (WPH): $z_{1*2} < \min(z_1, z_2)$.

A common index of heterosis, the potence ratio (Mather, 1949):

$$H_{PR} = \frac{z_{1*2} - \frac{z_1 + z_2}{2}}{\left| \frac{z_1 - z_2}{2} \right|},$$

directly gives the type of inheritance:

- If $H_{PR} > 1$, there is BPH.
- If $0 < H_{PR} \leq 1$, there is +MPH.
- If $H_{PR} = 0$, there is ADD.
- If $-1 \leq H_{PR} < 0$, there is –MPH.
- If $H_{PR} < -1$, there is WPH.

We frequently used in this study the index of best-parent heterosis:

$$H_{BP} = \frac{z_{1*2} - \max(z_1, z_2)}{\max(z_1, z_2)}.$$

- If $H_{BP} > 0$, there is BPH.
- If $H_{BP} \leq 0$, there can be +MPH, ADD, –MPH or WPH.

The mean of the positive H_{BP} values was noted $\overline{H_{BP}^+}$.

We also defined an index of worst-parent heterosis:

$$H_{WP} = \frac{z_{1*2} - \min(z_1, z_2)}{\min(z_1, z_2)}.$$

- If $H_{WP} < 0$, there is WPH.

- If $H_{WP} \geq 0$, there can be –MPH, ADD, +MPH or BPH.
The mean of the negative H_{WP} values was noted $\overline{H_{WP}^-}$.

Finally, we used in some instances the index of mid-parent heterosis:

$$H_{MP} = \frac{z_{1*2} - \frac{z_1 + z_2}{2}}{\frac{z_1 + z_2}{2}}.$$

- If $H_{MP} > 0$, there is BPH or +MPH.
- If $H_{MP} = 0$, there is ADD.
- If $H_{MP} < 0$, there is –MPH or WPH.

Note that in this context, various authors wrongly refer to heterosis as overdominance, complete dominance, partial dominance, and use various circumlocutions for negative heterosis. Strictly speaking, the words “dominance” and overdominance only apply to monogenic traits.

The Concave (Resp. Convex) Genotype-Phenotype Relationship Results in Positive (Resp. Negative) Heterosis

Consider $\tilde{z} = \tilde{f}(x_i) = f(x_1^\circ, \dots, x_i, \dots, x_n^\circ)$, defined as the function obtained by fixing $x_j = x_j^\circ, \forall j \neq i$, and letting x_i be variable. Assume that $\forall i, \tilde{z} = \tilde{f}(x_i)$ is strictly concave. Following the standard concavity argument, we can write:

$$\tilde{f}(tx_{i1} + (1-t)x_{i2}) > t\tilde{f}(x_{i1}) + (1-t)\tilde{f}(x_{i2}),$$

where $0 < t < 1$. The additivity of genotypic variables corresponds to $t = 1/2$ for all i , so we have:

$$\tilde{f}\left(\frac{x_{i1} + x_{i2}}{2}\right) > \frac{\tilde{f}(x_{i1}) + \tilde{f}(x_{i2})}{2}.$$

Hence

$$z_{1*2} > \frac{z_1 + z_2}{2}, \quad (1)$$

which means that +MPH or BPH necessarily holds, and that:

$$H_{BP} \in] -0.5, +\infty[.$$

Note that if the parents have the same phenotypic value ($z_1 = z_2$), it comes from Equation (1) that there is necessarily BPH.

If $\forall i, \tilde{z} = \tilde{f}(x_i)$ were strictly convex, Equation (1) would become:

$$z_{1*2} < \frac{z_1 + z_2}{2},$$

and there would be necessarily –MPH or WPH.

Application of These Theoretical Models to the Case of the Enzyme-Flux Relationship

Since Wright (1934), the enzyme-flux relationship is considered the archetype of the concave genotype-phenotype (GP)

relationship. Using a simple system of differential equations, he relied on the concavity of the GP curve to propose his famous model of dominance. Kacser and Burns (1981) refined his approach using a formalism in which the relationship between enzyme concentrations and the flux through a metabolic pathway is described by a heuristic equation which is at the heart of the metabolic control theory (Kacser and Burns, 1981; Heinrich and Schuster, 1996; Fell, 2007). Similarly, but more generally, Fiévet et al. (2006, 2010) expressed the relationship between the overall flux J ($M^{-1}.s^{-1}$) through a metabolic network and n enzyme concentrations $\{E_j\}_{j=1,\dots,n}$ in the following way:

$$J = f(E_1, \dots, E_n) = \frac{1}{\sum_{j=1}^n \frac{1}{XA_jE_j + Xd_jE_{tot}}}, \quad (2)$$

where E_j is the concentration of enzyme j , X is a positive constant, A_j and d_j are systemic parameters (unit $M^{-1}.s^{-1}$) that account for the kinetic behavior of enzyme E_j in the network, and E_{tot} is the total enzyme amount: $\sum_{j=1}^n E_j = E_{tot}$. Because in the standard case A_j and d_j are positive, the flux is an ascending function of E_j , for all j . We will consider that A_j and d_j are not genetically variable, so that variation in enzyme activity will only be due to the variability of the enzyme concentration E_j . Nevertheless formal developments with variable A_j 's would also be possible since A_j and E_j play a similar role in the relationship.

Flux heterosis

The variables and parameters of Equation (2) are all positive, and the function $J = f(E_1, \dots, E_n)$ has a negative partial second derivative with respect to E_j , for all j . Therefore, the function is strictly concave, and the previous developments apply. In particular, consider two individuals P1 and P2 that differ in their distribution of enzyme concentrations, respectively $\{E_{j1}\}_{j=1,\dots,n}$ and $\{E_{j2}\}_{j=1,\dots,n}$. Assuming additive inheritance of enzyme concentrations, the distribution of enzyme concentrations of the P1*P2 hybrid is:

$$\left\{ \frac{E_{j1} + E_{j2}}{2} \right\}_{j=1,\dots,n}$$

and its flux is:

$$J_{1*2} = \frac{1}{\sum_{j=1}^n \left(\frac{1}{XA_j \left(\frac{E_{j1} + E_{j2}}{2} \right) + Xd_j \left(\frac{E_{tot1} + E_{tot2}}{2} \right)} \right)}. \quad (3)$$

Due to concavity, we have for all j :

$$\tilde{f}\left(\frac{E_{j1} + E_{j2}}{2}\right) > \frac{\tilde{f}(E_{j1}) + \tilde{f}(E_{j2})}{2},$$

so

$$J_{1*2} > \frac{J_1 + J_2}{2}.$$

Thus, either +MPH or BPH are expected in this GP relationship, and if the parents have the same flux value ($J_1 = J_2$) there is necessarily BPH.

Conditions for best-parent heterosis

In the particular case where there is only one variable enzyme, BPH cannot occur, as can be shown from Equation (2). Otherwise, with $J_2 > J_1$, the condition for having BPH is $J_{1*2} > J_2$, i.e.,

$$\frac{1}{\sum_{j=1}^n \left(\frac{1}{XA_j \left(\frac{E_{j1} + E_{j2}}{2} \right) + Xd_j \left(\frac{E_{tot1} + E_{tot2}}{2} \right)} \right)} > \frac{1}{\sum_{j=1}^n \left(\frac{1}{XA_j E_{j2} + Xd_j E_{tot2}} \right)},$$

or

$$\sum_{j=1}^n \frac{1}{XA_j E_{j2} + Xd_j E_{tot2}} > \sum_{j=1}^n \frac{1}{XA_j \left(\frac{E_{j1} + E_{j2}}{2} \right) + Xd_j \left(\frac{E_{tot1} + E_{tot2}}{2} \right)}.$$

After rearrangements, we get:

$$\sum_{j=1}^n \left(a_j(E_{j1} - E_{j2}) + b_j(E_{tot1} - E_{tot2}) \right) > 0, \quad (4)$$

where

$$a_j = \frac{XA_j}{(XA_j E_{j2} + Xd_j E_{tot2}) \left(XA_j(E_{j1} + E_{j2}) + Xd_j(E_{tot1} + E_{tot2}) \right)}$$

and

$$b_j = \frac{Xd_j}{(XA_j E_{j2} + Xd_j E_{tot2}) \left(XA_j(E_{j1} + E_{j2}) + Xd_j(E_{tot1} + E_{tot2}) \right)}.$$

Note that $\forall j$, a_j and b_j are positive.

If the total enzyme concentration does not vary between individuals due to global cellular constraints of space and energy, we have $E_{tot1} = E_{tot2}$, and the condition for BPH is simplified:

$$\sum_{j=1}^n a_j(E_{j1} - E_{j2}) > 0. \quad (5)$$

In this case, enzyme concentrations do not vary independently, and this covariation will increase heterosis because the concavity of the flux-enzyme relationship is higher. This can be shown

by comparing the flux control coefficient (C_k^J) (Kacser and Burns, 1981) to the combined response coefficient (R_k^J) (Lion et al., 2004). Both coefficients measure the sensitivity of the flux to variations in enzyme concentration k , but the former applies when concentrations vary independently whereas the latter applies when there is a constraint on the total enzyme concentration. Following (Lion et al., 2004), we have:

$$R_k^J = C_k^J + E_k \sum_{j \neq k}^n C_j^J \frac{\alpha_{kj}}{E_j},$$

where $\alpha_{kj} = \frac{\partial E_j}{\partial E_k} = -\frac{E_j}{E_{tot}} \left(1 - \frac{E_k}{E_{tot}} \right)$ is the redistribution coefficient, i.e., the ratio of an infinitesimal change in concentration E_j to an infinitesimal change in concentration E_k . As $\forall j, C_j^J > 0$ and $\forall (k, j), \alpha_{kj} < 0$, we have $R_k^J < C_k^J$, which means that constraint on E_{tot} increases concavity.

Optimal distribution of enzyme concentrations

The optimal distribution of enzyme concentrations is the one that maximizes the flux through the network for a given E_{tot} value. The relevant unit of measure here is g.L^{-1} and not M.L^{-1} , because it is the total weight that matters in terms of energy and crowding in the cell. So we re-wrote Equation (2):

$$J = \frac{1}{\sum_{j=1}^n \frac{1}{XA'_j g E_j + Xd_j \sum_{j=1}^n \frac{g E_j}{M_j}}}, \quad (6)$$

where $A'_j = A_j/M_j$, with M_j the molecular weight of enzyme j , and where $g E_j$ is the concentration of enzyme j in g.L^{-1} .

The constraint is on $\sum_{j=1}^n g E_j = g E_{tot}$, which is different

from a constraint on $\sum_{j=1}^n \frac{g E_j}{M_j} = E_{tot}$. We derived the optimal

distribution of enzyme concentrations, $\{g E_i^*\}_{i=1, \dots, n}$, using the Lagrange multiplier method (see Fiévet et al., 2006 for details):

$$g E_i^* = \frac{g E_i^*}{g E_{tot}} = \frac{\frac{\sqrt{M_i/A_i}}{\sum_j 1/\sqrt{M_j A_j}} \left(1 + \sum_j \frac{d_j}{A_j} \right) - \frac{M_i d_i}{A_i}}{\sum_k \left(\frac{\sqrt{M_k/A_k}}{\sum_j 1/\sqrt{M_j A_j}} \left(1 + \sum_j \frac{d_j}{A_j} \right) - \frac{M_k d_k}{A_k} \right)}. \quad (7)$$

In Vitro and in Silico Genetics

In Vitro Crosses

The *in vitro* system based on the upstream part of glycolysis is described in Fiévet et al. (2006). It includes the following enzymes (see **Figure 2**): HK, hexokinase (E.C. 2.7.1.1); PGI, phosphoglucose isomerase (E.C. 5.3.1.9); PFK, phosphofructokinase (E.C. 2.7.1.11); FBA, fructose-1,6-bisphosphate aldolase (E.C. 4.1.2.13); TPI, triosephosphate

isomerase (E.C. 5.3.1.1); G3PDH, glycerol 3-phosphate dehydrogenase (E.C. 1.1.1.8) and CK, creatine phosphokinase (E.C. 2.7.3.2). All assays were carried out at 25°C in 50 mM Pipes buffer pH 7.5 containing 100 mM glucose, 100 mM KCl, 20 mM phosphocreatine, 3 mM NADH and 5 mM Mg-acetate. The reaction was started by the addition of 1 mM ATP. The ATP pool was kept constant via the regeneration reaction catalyzed by creatine phosphokinase. The high concentration of G3PDH (1 mM) allowed us to measure the flux through the pathway as the rate of NADH consumption at 390 nm, using an Uvikon 850 spectrophotometer.

In order to mimic genetic variability in enzyme activity, we varied the concentrations of PGI, PFK, FBA and TPI in test tubes to create “parents,” each parent being defined by a particular vector of concentrations. HK concentration

$$J = \frac{10^3}{\frac{1}{499.4 \frac{gE_{PGI}}{M_{PGI}}} + \frac{1}{115.5 \frac{gE_{PFK}}{M_{PFK}}} + \frac{1}{22.5 \frac{gE_{FBA}}{M_{FBA}}} + \frac{1}{22940 \frac{gE_{TPI}}{M_{TPI}} + 21.9 \sum_{j=1}^n \frac{gE_j}{M_j}}},$$

was fixed at 5.37 mg.L⁻¹. The total enzyme concentration of the four variable enzymes in the parental tubes was fixed at $E_{tot}^{\varphi} = 101.9$ mg.L⁻¹, a value chosen from the physiological concentrations estimated in yeast strain S288C (Fiévet et al., 2004) (PGI: 9.1 mg.L⁻¹, PFK: 10.4 mg.L⁻¹, FBA: 60.1 mg.L⁻¹ and TPI: 22.3 mg.L⁻¹). To have parental genotypes covering a large range of enzyme concentrations, we varied the concentration of FBA (the most abundant enzyme) from 0 to E_{tot}^{φ} , taking 10 values evenly distributed across this range (excluding of course the two extreme values). The proportions of the remaining enzymes (PGI, PFK and TPI) were drawn from beta distributions with shape parameter $\alpha = 1$ and scale parameter $\beta = (1 - e_i^{\varphi})/e_i^{\varphi}$, with e_i^{φ} the physiological proportion of enzyme i . We thus created 58 parental distributions.

“Hybrids” were produced by mixing 1:1 the content of the parental tubes, which corresponds to additivity of enzyme concentrations. In order to have hybrids derived from a large range of parental values, we defined four classes of parental flux values (< 5 μ M.s⁻¹, 5–8 μ M.s⁻¹, 8–11 μ M.s⁻¹ and > 11 μ M.s⁻¹), and we performed “crosses” within and between classes. We created 61 hybrids using 36 different parents (Table S1). For each of the 97 genotypes (61 hybrids + 36 parents), three flux measurements were performed (Fiévet et al., 2006), with the exception of two parents and four hybrids for which there were only two replicates (Table S2).

Two statistical tests were performed for every cross: (i) a Student's t -test to compare hybrid and mean parental fluxes; (ii) a one-way ANOVA with the genotype as factor (three levels: parent 1, parent 2 and hybrid) followed by a Tukey's test to classify observed inheritance as either BPH, +MPH, ADD, -MPH or WPH. For both, p -values were adjusted for multiple tests using the conservative Holm's method.

Computer Simulations of the *in Vitro* System

We simulated the *in vitro* system using the parameter values published by Fiévet et al. (2006), who estimated XA_j and

Xd_j by hyperbolic fitting of the titration curves obtained by varying the concentration of each enzyme in turn. The values were $XA_{PGI} = 499.4$ s⁻¹, $XA_{PFK} = 115.5$ s⁻¹, $XA_{FBA} = 22.5$ s⁻¹, $XA_{TPI} = 22940$ s⁻¹, $Xd_{PGI} = 0$, $Xd_{PFK} = 0$, $Xd_{FBA} = 0$ and $Xd_{TPI} = 21.9$ s⁻¹. Thus, for any set of E_j values, i.e., for any virtual genotype, the flux value could be computed from the equation:

$$J = \frac{1}{\frac{1}{499.4E_{PGI}} + \frac{1}{115.5E_{PFK}} + \frac{1}{22.5E_{FBA}} + \frac{1}{22940E_{TPI} + 21.9E_{tot}}},$$

where $E_{tot} = E_{PGI} + E_{PFK} + E_{FBA} + E_{TPI}$. Because enzyme concentrations in the simulations were in mg.L⁻¹ and flux is expressed in μ M.s⁻¹, we used the equation:

with the molecular masses $M_{PGI} = 61128$, $M_{PFK} = 36000$, $M_{FBA} = 39211$ and $M_{TPI} = 26700$.

We analyzed the effects of three factors on heterosis: (i) mean parental enzyme concentrations: either equidistributed or centered on their optima; (ii) constraint on total enzyme concentrations: either free E_{tot} or fixed E_{tot} ; (iii) changes in enzyme concentrations, by varying their coefficients of variation. For equidistributed enzyme concentrations, the means of the gamma distributions were $\mu = 25.475$ mg.L⁻¹ (i.e., $E_{tot}^{\varphi}/4$) whatever the enzyme. For the optimum-centered distributions, the means were $\mu = 15.768$ mg.L⁻¹, $\mu = 25.162$ mg.L⁻¹, $\mu = 59.497$ mg.L⁻¹ and $\mu = 1.473$ mg.L⁻¹ for PGI, PFK, FBA and TPI, respectively, computed from Equation (7). In all cases we varied the coefficients of variation from $c_v = 0.1$ to $c_v = 1.2$, within the range of observed c_v 's for these enzymes (Albertin et al., 2013). The shape and scale parameters of the gamma distributions were respectively $k = 1/c_v^2$ and $\theta = \mu c_v^2$.

To apply a strict constraint on the total amount of enzymes allocated to the system, we assumed that for each parent i the total enzyme amount, $\sum_j gE_{ji} = gE_{toti}$, was 101.9 mg.L⁻¹, namely the physiological total enzyme concentration. After drawing enzyme concentrations under the different conditions described above (equidistributed/centered on the optima for different c_v values), the concentration gE_{ki} of enzyme k in individual i was computed as:

$$gE_{ki} = 101.9 \frac{x_{ki}}{\sum_j x_{ji}}, \quad (8)$$

where x_{ki} was the result of the draw. Thus, the total enzyme concentration was 101.9 mg.L⁻¹ for every individual. It is worth noting that the constraint on E_{tot} changes the c_v 's in an inverse relation to enzyme concentrations: the c_v 's of the most (resp. the less) abundant enzymes will decrease (resp. increase) under constraint (see Appendix S1).

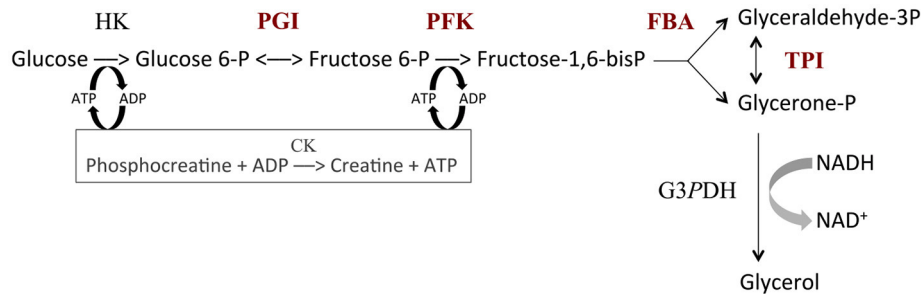


FIGURE 2 | The upstream part of glycolysis reconstructed *in vitro*. HK, hexokinase (E.C. 2.7.1.1); PGI, phosphoglucose isomerase (E.C. 5.3.1.9); PFK, phosphofructokinase (E.C. 2.7.1.11); FBA, fructose-1,6-bisphosphate aldolase (E.C. 4.1.2.13); TPI, triosephosphate isomerase (E.C. 5.3.1.1); G3PDH, glycerol 3-phosphate dehydrogenase (E.C. 1.1.1.8); CK, creatine phosphokinase (E.C. 2.7.3.2). Variable enzymes are in red. Reaction rate was measured from the decrease in NADH concentration.

For every simulation condition, we drew two series of 10,000 parents which were “crossed” to get 10,000 hybrids. Parent and hybrid fluxes were computed according to Equation (6), assuming additivity of enzyme concentrations.

Computer Simulations of the Glycolysis and Fermentation Network in Yeast

We modeled the glycolysis and fermentation network in yeast using a simplified topology in which there was one input flux, the rate of glucose consumption, and two output fluxes, the rate of glycerol production and the rate of transformation of pyruvate into acetaldehyde (**Figure 3**). The system of ordinary differential equations we used was derived from Conant & Wolfe’s model (Conant and Wolfe, 2007) obtained from the Biomodels database (<http://www.ebi.ac.uk/biomodels-main/>). This model is a slightly different version of a previously published model (Teusink et al., 2000). We modified this model in order to increase the range of enzyme concentrations that can lead to a steady state. First, we excluded the reactions producing trehalose and glycogen because, due to their constant rates, they caused incompatibility when the entry of glucose decreased. Second we suppressed mitochondrial shuttling, which is a composite reaction, because it is not associated with any enzyme. Thus, our model included 17 reactions and 22 metabolites, including 6 cofactors (ATP, ADP, AMP, NAD, NADH, and Fru-2,6-BP) (**Figure 3**). For all simulation conditions (see below), we solved the system of ordinary differential equations numerically using solvers (deSolve package) of the Runge-Kutta family (Hindmarsh, 1983) implemented within the R software tool (R Core Team, 2013). We considered that the steady state was reached when metabolite concentrations varied less than 10^{-4} between two successive time steps.

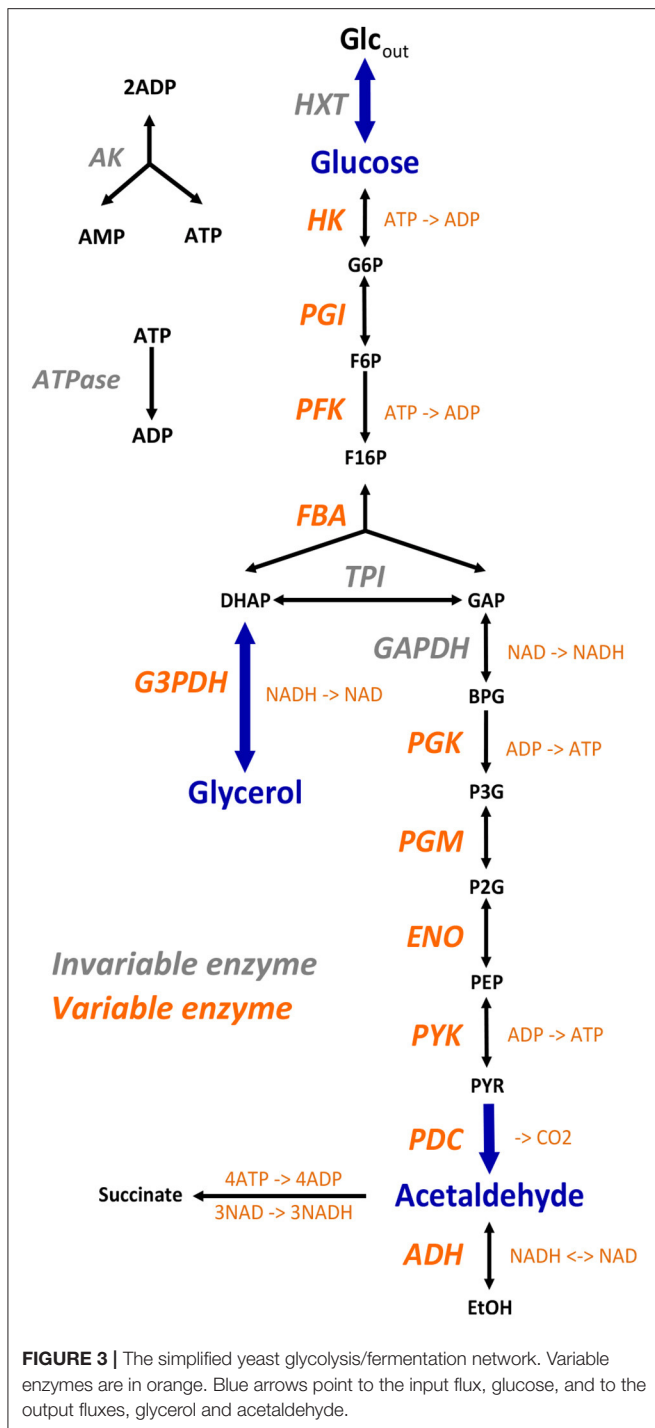
Conant & Wolfe’s model uses the maximal enzymatic rate $V_{\max}$, i.e., there is no explicit expression of enzyme concentration and catalytic constant k_{cat} . To simulate variation in enzyme concentrations we needed to know the k_{cat} values and molecular weights of the enzymes. We found most k_{cat} values in the Brenda database (<http://www.brenda-enzymes.info/>). When they were not known in yeast (in two cases), we used data from other organisms. Thus, we

got k_{cat} values for 11 enzymes (Table S3): HK, hexokinase (*S. cerevisiae*); PGI, phosphoglucose isomerase (*S. cerevisiae*); PFK, phosphofructokinase (*Schizosaccharomyces pombe*); FBA, fructose-1,6-bisphosphate aldolase (*S. cerevisiae*); PGK, 3-phosphoglycerate kinase (*S. cerevisiae*); PGM, phosphoglycerate mutase (*S. cerevisiae*); ENO, enolase (*S. cerevisiae*); PYK, pyruvate kinase (*S. cerevisiae*); PDC, pyruvate decarboxylase (*S. cerevisiae*); ADH, alcohol dehydrogenase (*Flavobacterium frigidimaris*) and G3PDH, glycerol 3-phosphate dehydrogenase (*Oryctolagus cuniculus*). For these enzymes, we determined the reference concentrations E_{ref} of the model from the equation:

$$E_{\text{ref}} = \frac{V_{\max} M_r}{k_{\text{cat}}},$$

where M_r is the molecular weight of the enzyme (Table S3).

As expected, the flux response to variation in concentration of each enzyme, with the other 10 concentrations being maintained at their reference values, displayed saturation curves in most cases. However, for certain enzymes the reference concentration was far or very far along the plateau, so that varying the concentration was almost without effect on the fluxes. Therefore, we tried to arbitrarily decrease the reference values of these enzymes, but as a consequence the rate of successful simulations decreased dramatically. In the end, we solved this problem by decreasing the 11 reference concentrations by the same factor, which was empirically fixed at 5. Thus, we simulated the genetic variability of parental enzyme concentrations by drawing values from gamma distributions with parameters $k = 1/c_v^2$ and $\theta = E_{\text{ref}} c_v^2 / 5$. Variances were set to have seven c_v ’s ranging from 0.1 to 0.7. With higher c_v ’s the steady state was frequently not reached, and c_v ’s smaller than 0.1 produced results similar to those obtained with $c_v = 0.1$. For each c_v value, we drew 10,000 pairs of parents. The enzyme concentrations of their hybrids were computed assuming additive inheritance. We considered two situations, free E_{tot} and fixed E_{tot} . For free E_{tot} , we used the concentration values drawn as described above. For fixed E_{tot} , the concentrations were scaled to make their sum equal to the sum of the reference enzyme concentrations (31833.97 mg), as



previously explained (Equation 8): ${}_gE_{ki} = 31833.97 \frac{x_{ki}}{\sum_j x_{ji}}$. We

checked that the input flux (glucose) was equal to the sum of the output fluxes (glycerol and acetaldehyde), indicating that there was no “leak” in the system.

For each parent-hybrid triplet we computed the fluxes of glucose, glycerol and acetaldehyde at the steady state. There were

two possible sources of bias on the *a posteriori* c_v 's. First, if one member of the triplet did not reach a steady state, the triplet was excluded. Thus, to get 10,000 simulations, $\approx 10,060$ to $\approx 36,200$ simulations were necessary, depending on the c_v and the presence of constraint on E_{tot} . Second, even if a steady state was reached, the glucose flux could be null, and we obviously eliminated the corresponding triplets. An analysis of the resulting biases is presented in Appendix S2. Overall, the differences between *a priori* and *a posteriori* c_v values were insignificant or moderate when initial c_v 's were small to medium, and were significant only for certain enzymes with high initial c_v 's. In any case, these differences maintained the wide range of c_v 's, making it possible to study the effects of variability in enzyme concentrations on heterosis.

Flux Differences and Enzymatic Distances Between Parents

The relationship between parental flux values and heterosis was analyzed using the absolute value of the flux difference between parents i and i' :

$$D_{flux_{ii'}} = |J_i - J_{i'}|,$$

where J_i (resp. $J_{i'}$) is the mean flux of parent i (resp. i').

The relationship between parental enzyme concentrations and heterosis was analyzed using the weighted normalized Euclidean distance between parents i and i' :

$$D_{enz_{ii'}} = \sqrt{\frac{\sum_{j=1}^n \left(\frac{{}_gE_{ji} - {}_gE_{ji'}}{{}_gE_{ji} + {}_gE_{ji'}} \right)^2}{n}},$$

where n is the number of enzymes. With normalization, we have $0 \leq D_{enz_{ii'}} \leq 1$.

To examine the relationships between the heterosis indices and the distances, we used – when applicable – the Pearson correlation coefficient (r) and/or the coefficient of determination (R^2). Otherwise we used dedicated graphical representation methods (see Results section).

RESULTS

In order to analyze the way the genotype-phenotype (GP) relationship shapes heterosis, we used theoretical developments, *in vitro* genetics and computer simulations of the glycolytic/fermentation flux.

A Geometrical Approach: The Concavity of the GP Relationship Results in Heterosis

As an archetype of the concave GP relationship, we first chose a hyperbolic function relating the flux J_i of individual i to enzyme concentrations in a n -enzyme metabolic system:

$$J_i = \frac{1}{\sum_{j=1}^n \frac{1}{XA_j E_{ji} + Xd_j E_{toti}}}, \quad (9)$$

where X is a positive constant, A_j and d_j are systemic positive parameters accounting for the kinetic behavior of enzyme E_j in the network, and E_{ji} and E_{tot_i} are respectively the concentration of enzyme E_j and the total enzyme concentration in individual i (Fiévet et al., 2006, 2010; Fell, 2007). Because we assumed in this study that there was additivity of enzyme concentrations, the flux of the hybrid of parents P1 and P2, with fluxes J_1 and J_2 respectively, was written:

$$J_{1*2} = \frac{1}{\sum_{j=1}^n \frac{1}{XA_j\bar{E}_j + Xd_j\bar{E}_{tot}}}. \quad (10)$$

where $\bar{E}_j = (E_{j1} + E_{j2})/2$ and $\bar{E}_{tot} = (E_{tot1} + E_{tot2})/2$. Since the relationship between enzyme concentrations and flux is a rectangular multivariate hyperbolic function, the concavity of the surface results in positive heterosis for the flux ($J_{1*2} > \frac{J_1 + J_2}{2}$), i.e., either positive Mid-Parent Heterosis (+MPH) or Best-Parent Heterosis (BPH) (see Figure S1 for definitions of the types of heterosis). This has been shown analytically (see Theoretical background), and is illustrated geometrically on the two-dimensional hyperbolic flux response surface obtained with two variable enzymes (Figure 4A). Interestingly, it is possible to show that there is BPH, i.e., $J_{1*2} > J_2$ (for $J_2 > J_1$), whenever:

$$\sum_{j=1}^n \left(a_j(E_{j1} - E_{j2}) + b_j(E_{tot1} - E_{tot2}) \right) > 0, \quad (11)$$

where a_j and b_j are positive terms (see Theoretical background). This equation predicts that heterosis will be observed only if a sufficient number of enzymes have a lower concentration in the parent displaying the highest flux than in the other parent, making $(E_{j1} - E_{j2})$ positive, and this effect is strengthened if E_{tot} is smaller in the parent that has the highest flux. These conditions are more likely to arise when parental fluxes are similar (Figure 4A). In other words, a condition for observing BPH is that parents display contrasted distributions of enzyme concentrations, i.e., with complementary “high” and “low” alleles.

In the case of constrained resources, i.e., if the total enzyme amount allocated to the system is limited, which is biologically realistic, the flux-enzyme relationship is no longer a multivariate hyperbolic function but exhibits a hump-shaped response surface because enzyme concentrations are negatively correlated (Figure 4B). Increasing the concentration of certain enzymes leads to decreasing the concentration of other enzymes (Heinrich and Schuster, 1998; Lion et al., 2004; Malmstroem et al., 2009). The constraint on E_{tot} increases the concavity of the surface, which reinforces the incidence of heterosis (see Theoretical background). If E_{tot} is the same in the two parents, the condition for BPH is simply written as:

$$\sum_{j=1}^n a_j(E_{j1} - E_{j2}) > 0.$$

Note that when E_{tot} is constant, BPH cannot be observed if a parent has optimal enzyme concentrations, because no offspring can exceed the maximal flux value.

In Vitro and in Silico Heterosis in a Four-Enzyme Pathway

In order to assess these predictions, we performed *in vitro* genetic experiments. We reconstructed the first part of glycolysis (Figure 2), and created 61 “hybrid” tubes by mixing 1:1 the content of “parental” tubes that differed in their distribution of the concentrations of four enzymes, PGI, PFK, FBA and TPI, with total enzyme concentration being constant (see Materials and Methods).

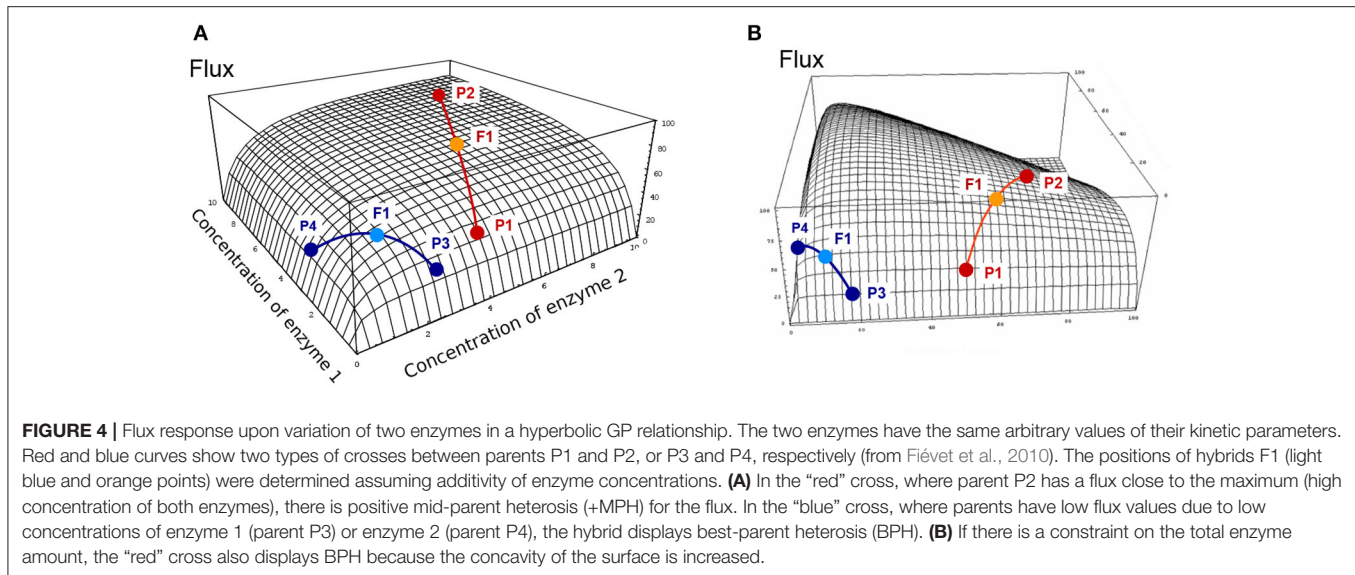
Inheritance was clearly biased toward high hybrid values (Figure 5). Among the 61 hybrids, 27 ($\approx 44.3\%$) exhibited a significantly higher flux value than the mid-parental value: seven ($\approx 11.5\%$) displayed BPH and 20 ($\approx 32.8\%$) displayed +MPH (p -values < 0.05). The maximum BPH value, quantified with the index $H_{BP} = \frac{J_{1*2} - \max(J_1, J_2)}{\max(J_1, J_2)}$, was $H_{BP} = 0.37$, i.e., the hybrid displayed a flux that was 37% higher than the best parental flux. By contrast, only eight hybrids had a flux value that was significantly lower than their mid-parental value (negative Mid-Parent Heterosis, or -MPH), with no case found of significant Worst-Parent Heterosis (WPH) (Figure 5 and Table S1).

Examining enzyme concentrations in the parents confirmed that high values of BPH were likely to occur when parental distributions were contrasted and when neither parent had enzyme concentrations close to the optimal distribution (Figures 5D,E). However, as shown in Figure 6A, this relationship is not straightforward, with large distances being necessary but not sufficient to get high H_{BP} values, which resulted in a non-significant correlation between H_{BP} and D_{enz} , the Euclidean distance between parental enzyme concentrations.

To test the prediction that a cross between parents with similar fluxes is more likely to give heterosis than a cross between parents with contrasted fluxes, we computed the correlation coefficient between H_{BP} and D_{flux} , the difference between parental fluxes ($|J_1 - J_2|$), and found a highly negative significant value ($r = -0.33, p < 0.01$) (Figure 6B).

Interestingly, even though H_{BP} and D_{enz} were not correlated, the heterosis value H_{reg} , computed from the equation of the multiple linear regression performed with both D_{enz} and D_{flux} as predictor variables, was better correlated with H_{BP} than D_{flux} alone: $r = 0.46$ ($p < 1.8 \cdot 10^{-4}$) (Figure 6C).

Finally, in order to know if the instances of heterosis we observed *in vitro* were consistent with our heuristic model of flux-enzyme relationship, we computed from Equations (9) and (10) the theoretical fluxes of the 61 hybrids and their parents, using the values of enzyme concentrations of the parental tubes and the values of parameters XA_j and Xd_j estimated in an independent experiment (see Materials and Methods). Then, for each virtual cross, we calculated the heterosis that would be expected from the model, H_{BPmod} . As shown in Figure 6D, the correlation between H_{BPmod} and H_{BP} is very highly significant ($r = 0.77, p < 2.5 \cdot 10^{-13}$).



As Equations (9) and (10) resulted in quite reliable predictions of flux heterosis, we used them to simulate a large series of crosses. The parental enzyme concentrations were drawn to be centered on their optimum or equidistributed, and their coefficients of variation (c_v) varied from 0.1 to 1.2. In every case we considered two situations, free and fixed E_{tot} . For every simulation condition, 10,000 crosses were performed and the parental and hybrid values were computed.

The percentage of BPH and the mean of positive H_{BP} values (H_{BP}^+) were compared over the range of c_v 's. **Figure 7** shows that: (i) in all conditions, the values of these two variables increased, usually to a large extent, when enzyme variability increased; (ii) by far the highest % BPH (about 50%) was observed when enzyme concentrations were constrained (fixed E_{tot}) and centered on their optimum (**Figure 7A**); (iii) H_{BP}^+ was highly dependent on the c_v values (varying from ≈ 0 to ≈ 1.4), but was barely affected by the other simulation conditions (**Figure 7B**).

The typical influence of D_{flux} on H_{BP} is illustrated **Figure 8** for $c_v = 0.6$. BPH was observed when the parental fluxes were close to each other, as the hybrids displaying BPH were distributed around the diagonal of the parental flux space. This pattern is consistent with the negative relationship between D_{flux} and H_{BP} (**Figure 9A**). The highest D_{flux} never resulted in BPH, while small D_{flux} could result in high H_{BP} values. The weakest correlation between D_{flux} and H_{BP} was observed when both enzyme concentrations were centered on their optimum and E_{tot} was fixed ($r = -0.35$ instead of $r > -0.66$ in other conditions).

The graphs showing the relationship between D_{enz} and H_{BP} revealed a typical triangular relationship, regardless of the simulation conditions: BPH was never observed with the smallest distances, and large distances were necessary but not sufficient to get high H_{BP} values (**Figure 9B**).

To assess the joint influence of D_{flux} and D_{enz} on heterosis, we divided the crosses into 16 classes according to a 4×4 grid of D_{flux} and D_{enz} values, and computed for each class the % BPH and H_{BP}^+ .

In all conditions, the highest values of these two variables were observed with both small D_{flux} and large D_{enz} (**Figures 9C,D**). As expected, constraint on E_{tot} resulted in more heterosis, with almost 100% BPH when: (i) enzyme concentrations were centered on their optimum; (ii) D_{enz} values were maximal; (iii) D_{flux} values were minimal (**Figure 9C**). In these conditions, H_{BP}^+ increased to 0.37, i.e., hybrids had on average a flux that was 37% higher than the best parental flux (**Figure 9D**).

In Silico Heterosis in the Yeast Glycolysis/Fermentation Network

To simulate heterosis in a larger system, we modeled the glycolysis and fermentation network in yeast using a simplified topology in which there was one input flux, the rate of glucose consumption, and two output fluxes, the production rates of glycerol and acetaldehyde (**Figure 3**). We used a system of differential equations derived from Conant and Wolfe (2007) and varied the concentration of 11 enzymes.

We first examined how varying the concentration of one enzyme at a time, the other concentrations being maintained at their reference values (see Material and methods), affected flux variation. When E_{tot} was free, the enzyme-flux relationships were concave with a horizontal asymptote in most cases, although a slight convexity was visible in two instances (**Figure S2**): (i) sigmoidal curves were observed upon variation of HK, an allosteric enzyme that is by far the most abundant in the system (**Table S3**); (ii) due to the balance between the acetaldehyde and glycerol branches, increasing the concentration of an enzyme in one branch convexly decreases the flux of the other branch (except for enzyme at low concentrations). The decrease of the glucose flux upon G3PDH increase is likely to be due to the related reduction of the acetaldehyde flux, which dampens ATP production and in turn reduces the glucose flux.

When E_{tot} was fixed, these observations remain qualitatively valid, however we no longer found horizontal asymptotes:



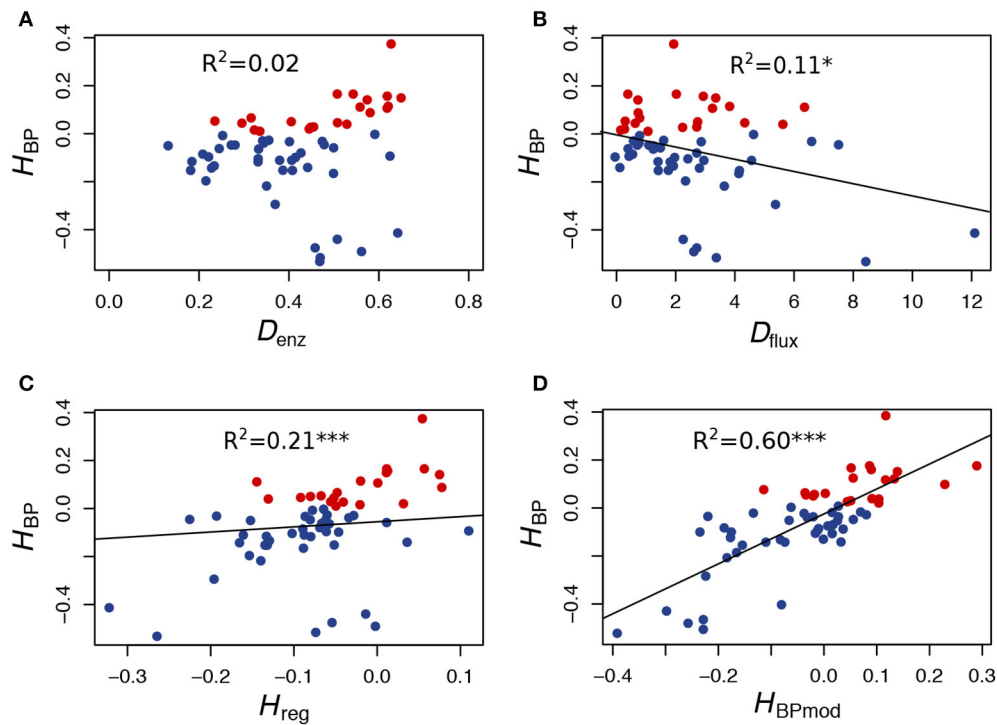


FIGURE 6 | *In vitro* heterosis predictors. **(A)** Relationship between H_{BP} and D_{enz} , the Euclidean distance between parents computed from enzyme concentrations ($r = 0.15$, $p < 0.25$). **(B)** Relationship between H_{BP} and D_{flux} , the flux difference between parents ($r = -0.33$, $p < 0.01$). **(C)** Relationship between H_{BP} and H_{reg} , the heterosis value computed from the equation of the multiple linear regression performed with D_{enz} and D_{flux} as predictor variables ($r = 0.46$, $p < 0.001$). **(D)** Relationship between H_{BP} and H_{BPmod} , the expected heterosis value computed from the model (see text) ($r = 0.77$, $p < 2.5 \cdot 10^{-13}$). Red points correspond to positive H_{BP} , i.e., to BPH.

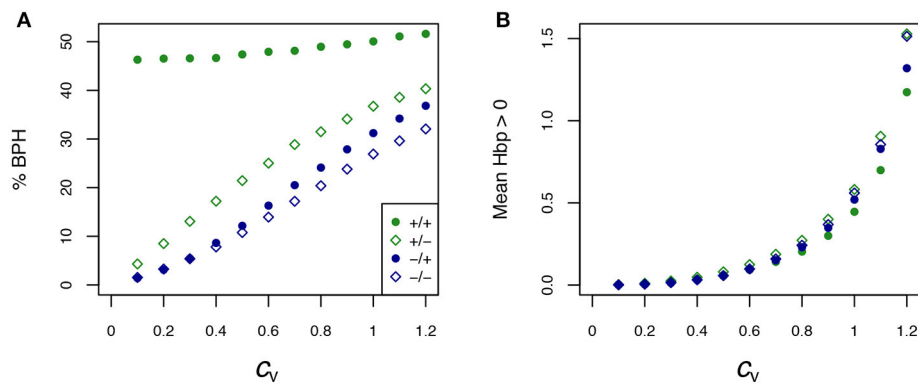


FIGURE 7 | Effect of the coefficient of variation (c_v) of enzyme concentrations on heterosis. **(A)** Percentage of BPH; **(B)** Mean of the positive H_{BP} values ($\overline{H_{BP}^+}$) over a range of c_v 's of parental enzyme concentrations, from 0.1 to 1.2. $-/-$: equidistributed mean enzyme concentrations and free E_{tot} , $-/+$: equidistributed and fixed E_{tot} , $+/-$: mean enzyme concentrations centered on their optimum and free E_{tot} , $+/+$: optimum centered means and fixed E_{tot} .

due to this constraint, flux values decreased when enzyme concentrations were high (Figure S3). This effect was marked for the most abundant enzymes (particularly HK and PYK) but was imperceptible for less abundant ones (e.g., PFK). Note that in all situations the three fluxes are null when enzyme concentrations are null, even for enzymes downstream of the bifurcation. Cofactor availability probably explains this observation.

We then simulated 10,000 crosses between parents that differed in their concentrations of the 11 variable enzymes. Parental concentrations were drawn from gamma distributions, the means of which were the reference values. Variances were chosen within the range of seven c_v 's, from 0.1 to 0.7, with free or fixed E_{tot} . Hybrid fluxes and inheritance were computed for each cross (see Materials and Methods). Unlike the previous

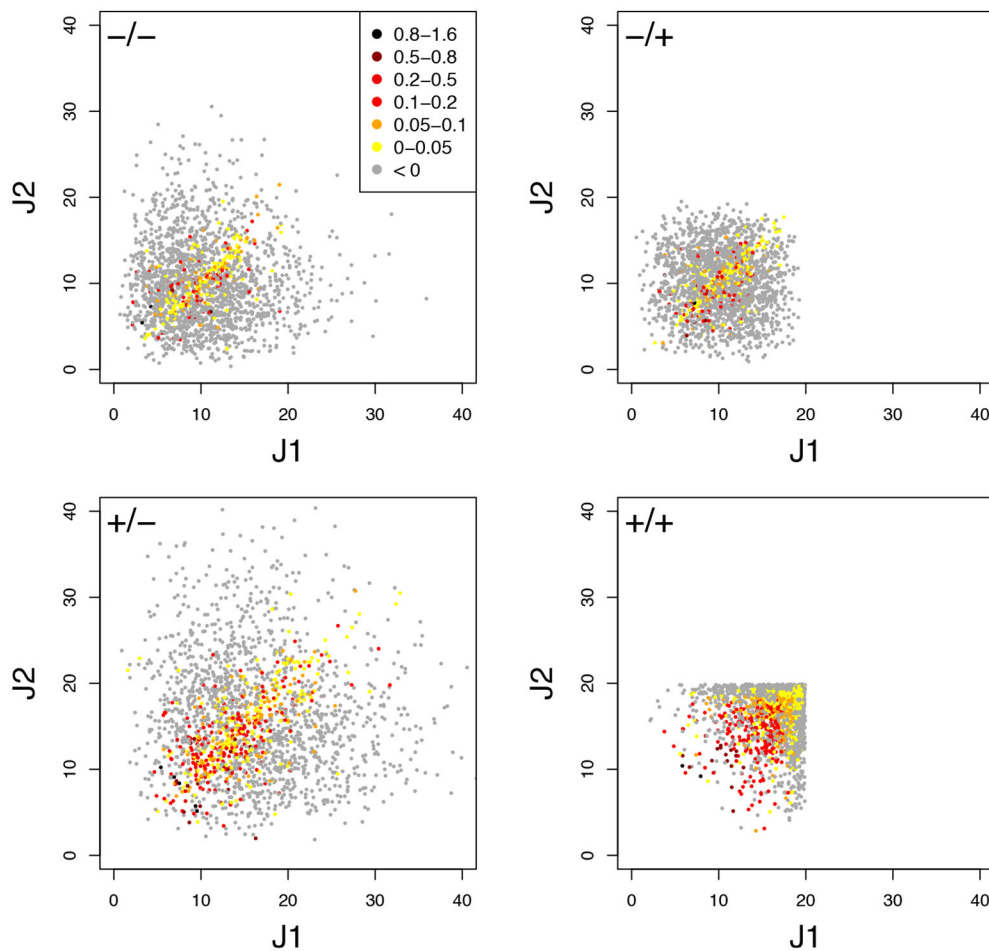


FIGURE 8 | Relationship between parental fluxes J_1 and J_2 and hybrid fluxes under four simulation conditions. Each point corresponds to a hybrid, with warm colors when there is BPH ($H_{BP} > 0$). $+/+$, $+/-$, $-/+$, and $-/-$ have the same meaning as in **Figure 7**. Note that when E_{tot} is fixed, the flux value is limited. ($c_v = 0.6$).

four-enzyme model, where the enzyme-flux relationship was concave by design and could only produce BPH or positive MPH, we observed the four possible types of inheritance, BPH, +MPH, -MPH and WPH. For the glucose and acetaldehyde fluxes, inheritance was very biased toward positive heterosis whatever the c_v , since the sum of the percentages of +MPH and BPH varied from 69.4 to 96.9%, and there was no or very few cases of WPH. The highest percentage of WPH, observed for acetaldehyde when $c_v = 0.7$ with free E_{tot} , was only 0.09%, with a very small mean of the negative H_{WP} values ($H_{WP}^- = -0.024$) (**Tables 1, 2** and **Figure 10**). The one-step glycerol branch, which is convexly related to the six enzymes of the acetaldehyde branch, had higher percentages of negative heterosis (from $\approx 65\%$ when $c_v = 0.1$ to $\approx 43\%$ when $c_v = 0.7$ with fixed E_{tot}), but WPH remained quite low: 1.23 and 3.02% for free and fixed E_{tot} , respectively, with $H_{WP}^- \approx -0.04$ in both cases ($c_v = 0.7$). For all fluxes, constraint on E_{tot} decreased the percentage of -MPH in favor of BPH and/or +MPH, depending on the c_v value (**Tables 1, 2** and **Figure 10**).

We hypothesized that negative heterosis can occur only when the enzyme-flux relationship is convex, as exemplified for negative dominance (**Figure 1A**). To verify this, we ran

simulations where HK concentration was maintained at its reference value and/or the glycerol branch was deleted. When both sources of convexity were removed, no case of negative heterosis was observed (**Table S4** and **Figure S4**).

As previously, both % BPH and H_{BP}^+ increased as the c_v increased (**Tables 1–3**, **Figures 11A,B**). This effect was larger for glucose and acetaldehyde than for glycerol. For instance, for glucose the % BPH varied from 3.76 to 35.12% with free E_{tot} and from 11.96 to 42.46% with fixed E_{tot} , while for glycerol it varied from 2.92 to 18.84% with free E_{tot} and from 3.71 to 24.10% with fixed E_{tot} . For the three fluxes, constraint on E_{tot} increased the % BPH, as expected, but decreased H_{BP}^+ . The latter observation is explained by the fact that this constraint decreases the variance of the concentrations, which in turn decreases the flux variance and hence both H_{BP} variance and H_{BP}^+ (see Appendix S1 and Appendix S2).

Whatever the c_v , BPH was usually observed when parental fluxes were similar, with heterotic hybrids roughly distributed around the diagonal of the parental flux space, and this was valid for the three fluxes (**Figure S5**). The correlations between H_{BP} and D_{flux} were consistent with these observations. They were negative

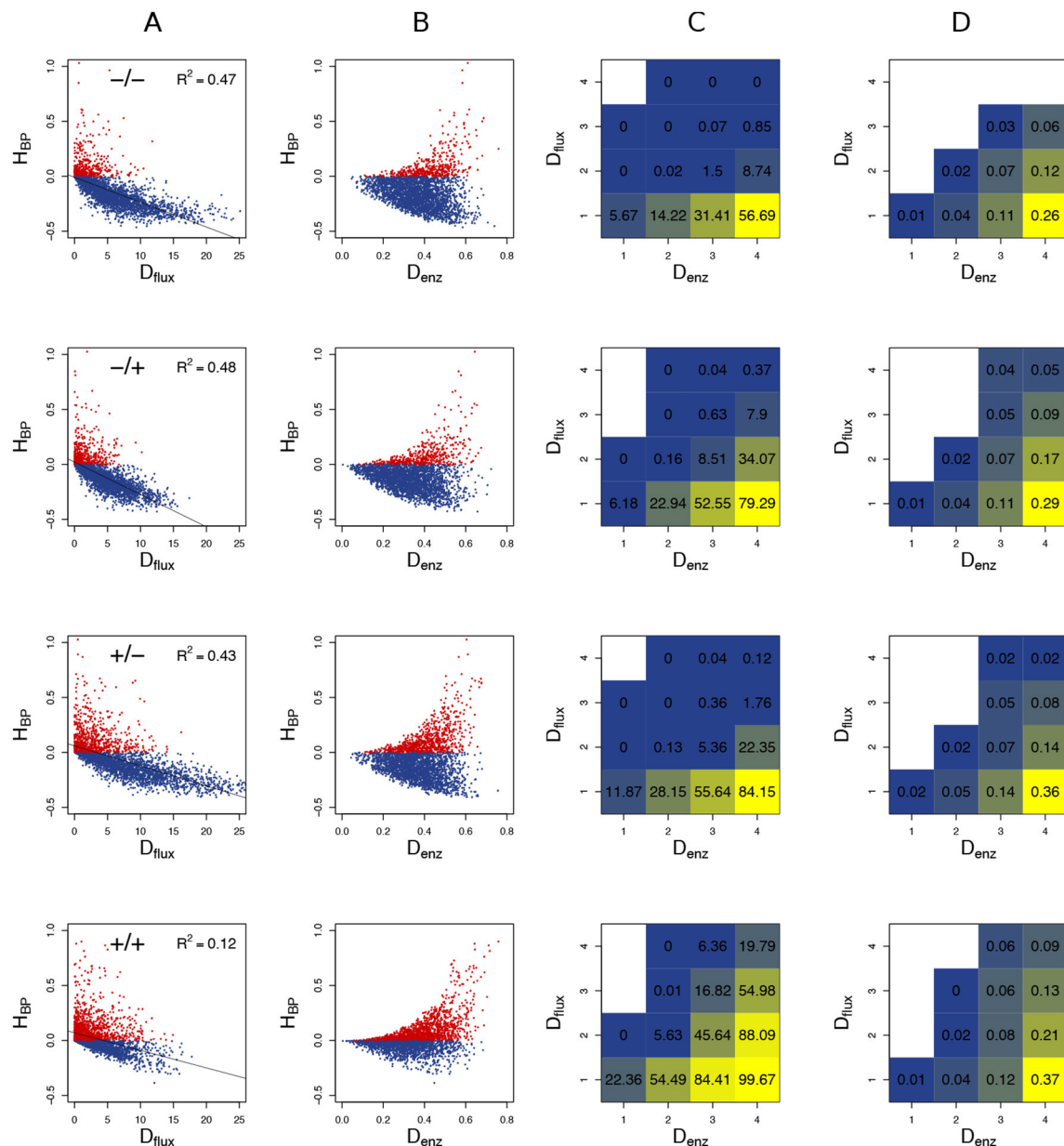


FIGURE 9 | Heterosis predictors. Relationship between H_{BP} and D_{flux} (A), and between H_{BP} and D_{enz} (B) when $c_v = 0.6$. Red points: positive H_{BP} values. (C) Square grid of percentages of BPH as a function of D_{enz} (x-axis) and D_{flux} (y-axis). Colors range from blue (minimum % BPH) to yellow (maximum % BPH). Empty squares are white. (D) Same as (C) with H_{BP}^+ values. The additional white squares correspond to cases where there was no BPH (0 in C). $-/-$, $-/+$, $+/-$, and $+/+$ have the same meaning as in Figure 7.

and highly significant, although slightly smaller when E_{tot} was fixed ($r = -0.60, -0.69$ and -0.61 for glucose, glycerol and acetaldehyde, resp.) than when E_{tot} was free ($r = -0.55, -0.59$ and -0.54 , resp.) (Figures 12A,C).

The triangular relationship between H_{BP} and D_{enz} was slightly less apparent with free E_{tot} than in the four-enzyme model, but was obvious with fixed E_{tot} . Again high H_{BP} values were never observed with small distances, and the highest H_{BP} values could only be observed with medium or large distances (Figures 12B,D).

As previously, we assessed the joint influence of D_{flux} and D_{enz} on % BPH and H_{BP}^+ using 4×4 grids (Figure 13, for $c_v = 0.4$). Results were very similar: both variables were positively related to D_{enz} and negatively related to D_{flux} , whatever the flux. High D_{enz} associated with small D_{flux} could result in about 50% BPH and $H_{BP}^+ \approx 0.25$ when E_{tot} was free and almost 100% BPH and $H_{BP}^+ \approx 0.40$ when E_{tot} was fixed. The positive effect of the constraint on H_{BP}^+ is not in contradiction with the above-mentioned argument: H_{BP}^+ was computed here for each square of the grid, and its

TABLE 1 | Percentage of occurrence of the four types of inheritance for the glucose, glycerol and acetaldehyde fluxes over the range of c_V values when E_{tot} is free.

c_V		0.1	0.2	0.3	0.4	0.5	0.6	0.7
Glucose	BPH	3.76	13.80	19.77	23.25	25.44	29.97	35.14
	+MPH	65.67	59.44	55.89	56.12	56.04	54.09	50.35
	−MPH	30.57	26.76	24.34	20.64	18.51	15.94	14.52
	WPH	0.00	0.00	0.00	0.00	0.01	0.04	0.00
Glycerol	BPH	2.92	10.84	14.92	16.94	16.99	18.39	18.85
	+MPH	31.77	33.61	33.49	35.12	38.38	37.41	38.27
	−MPH	64.83	54.85	50.66	47.01	43.40	43.14	41.65
	WPH	0.48	0.70	0.93	0.93	1.23	1.06	1.23
Acetaldehyde	BPH	4.05	14.63	21.23	24.88	28.05	32.71	37.76
	+MPH	68.10	60.42	56.15	55.79	53.92	52.49	48.45
	−MPH	27.86	24.93	22.62	19.32	18.03	14.80	13.69
	WPH	0.00	0.01	0.00	0.01	0.00	0.00	0.09

TABLE 2 | Percentage of occurrence of the four types of inheritance for the glucose, glycerol and acetaldehyde fluxes over the range of c_V values when E_{tot} is fixed.

c_V		0.1	0.2	0.3	0.4	0.5	0.6	0.7
Glucose	BPH	11.96	25.86	32.89	35.58	37.57	39.14	42.46
	+MPH	84.92	70.44	62.00	58.60	54.89	52.90	49.29
	−MPH	3.12	3.70	5.11	5.82	7.53	7.95	8.25
	WPH	0.00	0.00	0.00	0.00	0.01	0.01	0.00
Glycerol	BPH	3.71	13.64	18.72	20.07	21.57	23.19	24.10
	+MPH	31.03	30.33	31.54	33.10	34.36	34.46	35.78
	−MPH	64.16	53.74	46.95	43.98	41.12	39.51	37.09
	WPH	1.10	2.28	2.79	2.84	2.94	2.84	3.02
Acetaldehyde	BPH	12.61	26.93	33.77	36.83	38.91	41.01	44.40
	+MPH	83.23	68.28	60.07	56.40	52.63	50.62	46.71
	−MPH	4.16	4.80	6.16	6.77	8.46	8.35	8.89
	WPH	0.00	0.00	0.00	0.00	0.00	0.01	0.00

distribution was more uneven under constraint. On average $\overline{H_{\text{BP}}^+}$ was lower under constraint, but reached higher values when there was both small D_{flux} and high D_{enz} .

Figure S6 shows that contrasted distributions of enzyme concentrations can lead to BPH whereas close distributions favor additivity. Indeed, the cross that is the most heterotic for both glucose and acetaldehyde ($H_{\text{BP}} = 4.23$ and $H_{\text{BP}} = 5.55$, resp.) had divergent enzyme concentrations ($D_{\text{enz}} = 0.68$) (Figures S6A,B); this was also observed for the most heterotic cross for glycerol ($D_{\text{enz}} = 0.66$) (Figures S6C,D). These enzymatic distances are close to the highest distance observed ($D_{\text{enz}} = 0.71$), which also corresponded to large heterosis for glucose and acetaldehyde ($H_{\text{BP}} = 0.69$ and $H_{\text{BP}} = 1.09$, resp.) (Figures S6E,F). By contrast, the cross with the minimum enzymatic distance ($D_{\text{enz}} = 0.16$) exhibited additivity or near additivity for the three fluxes ($H_{\text{MP}} = 0$, $H_{\text{MP}} = 0.02$ and $H_{\text{MP}} = -0.01$ for glucose, glycerol and acetaldehyde, resp.) (Figures S6G,H).

The geometric models of **Figure 14** and **Figure S7** show a relationship between D_{enz} , D_{flux} and H_{BP} that is fully consistent

with all our observations. For a given value of D_{enz} , H_{BP} can be positive or negative depending on the position of the parents in the enzyme concentration space, with quite complex relationships between these two variables. BPH is observed when parents have complementary enzyme concentrations, with the highest H_{BP} values observed when D_{flux} is small. Of note, D_{flux} and H_{MP} are roughly positively related, because the reference for this index is the mid-parent and not the best parent (**Figures 14E,F**). Regarding the constraint on E_{tot} , comparison of **Figure 14A** and **Figure S8** shows how greater surface curvature increases the incidence of heterosis.

DISCUSSION

The So-Called Mystery of Heterosis

Nearly all papers on heterosis mention in their introduction that it is an elusive, enigmatic or even magical phenomenon, whose molecular (or physiological, or biochemical) bases are not known or not well understood, and which would be the “lasting mystery

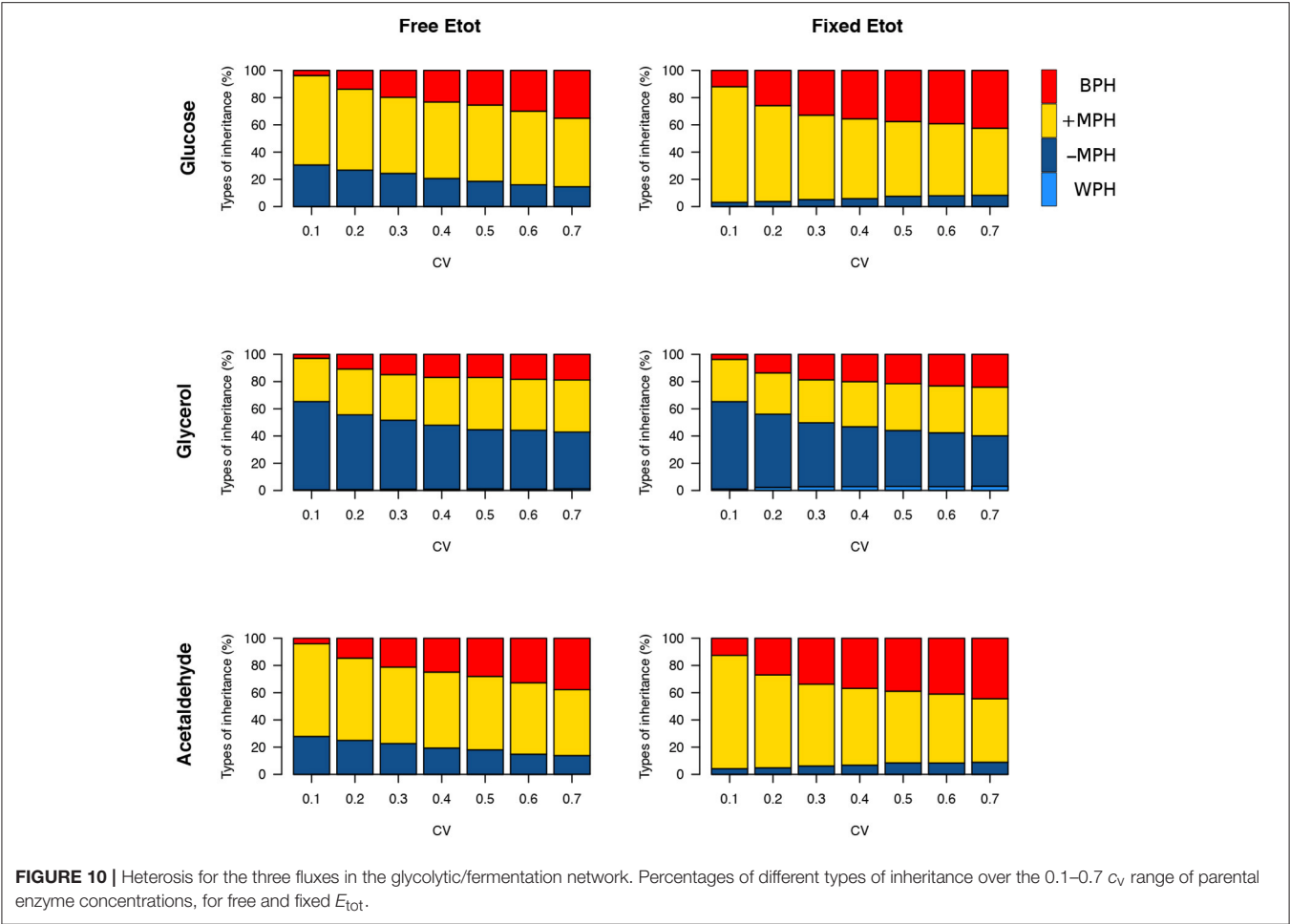


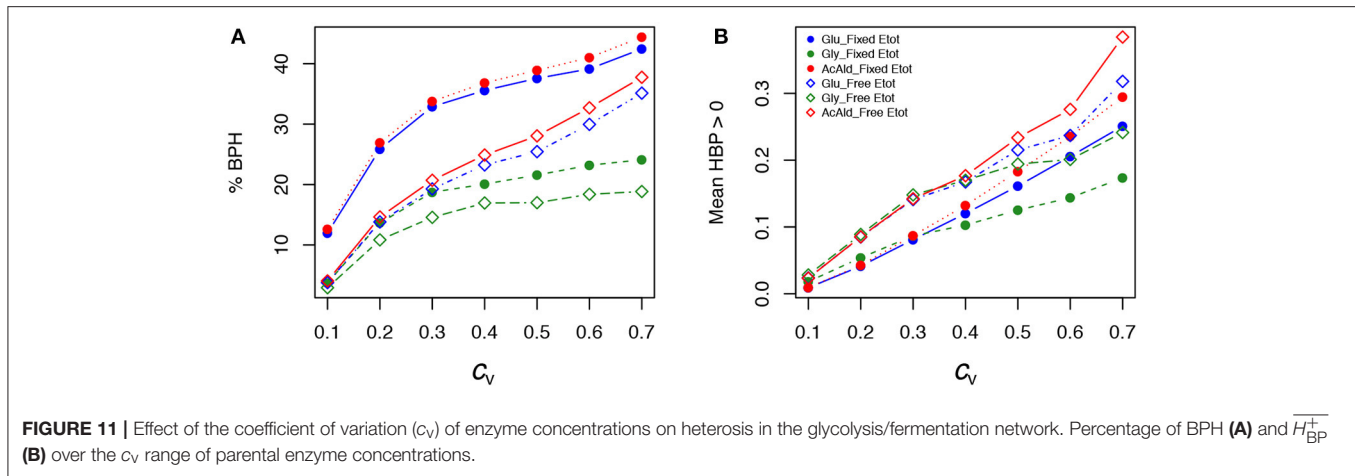
FIGURE 10 | Heterosis for the three fluxes in the glycolytic/fermentation network. Percentages of different types of inheritance over the 0.1–0.7 c_v range of parental enzyme concentrations, for free and fixed E_{Tot} .

TABLE 3 | Mean of positive H_{BP} values (H_{BP}^+) for the glucose, glycerol and acetaldehyde fluxes over the range of c_v values.

c_v		0.1	0.2	0.3	0.4	0.5	0.6	0.7
Free E_{Tot}	Glucose	0.02	0.08	0.14	0.17	0.22	0.24	0.32
	Glycerol	0.03	0.09	0.15	0.17	0.19	0.20	0.24
	Acetaldehyde	0.02	0.09	0.14	0.18	0.23	0.28	0.38
Fixed E_{Tot}	Glucose	0.01	0.04	0.08	0.12	0.16	0.21	0.25
	Glycerol	0.02	0.05	0.09	0.10	0.13	0.14	0.17
	Acetaldehyde	0.01	0.04	0.09	0.13	0.18	0.24	0.29

in biology.” Yet, this “fascination with the idea that standard genetic models are not sufficient to explain heterosis” (Kaeppler, 2011) is quite surprising given the literature on this phenomenon. First, regardless of any genetic or molecular basis, it was recognized early on (Richey, 1942; Grafius, 1961) that multiplicative traits, i.e., traits determined as the product of two or more components, display commonly predictable levels of heterosis. For instance, heterosis for plant height in bean is quite well explained by multiplying internode length by internode number (Coyne, 1965); similarly, yield heterosis in barley is predicted by the product of ears/plant, kernels/ear and average kernel weight (Grafius, 1959), etc. It is worth noting

that this cause of heterosis does not even require non-additive inheritance of the components of the trait (reviewed in Schnell and Cockerham, 1992). Second, heterosis has been explained at the genetic and/or molecular level for various traits in various species. Paradoxically, the best understood cases of heterosis are certainly the least common, since they involve only one locus, which makes them easier to study. Here, heterosis is due to overdominance, i.e., the inherent superiority of the heterozygote over both homozygotes (East, 1936; Hull, 1946). Relative to the innumerable examples of complete or partial dominance, well-attested overdominance remains quite rare, with probably less than 20 cases described

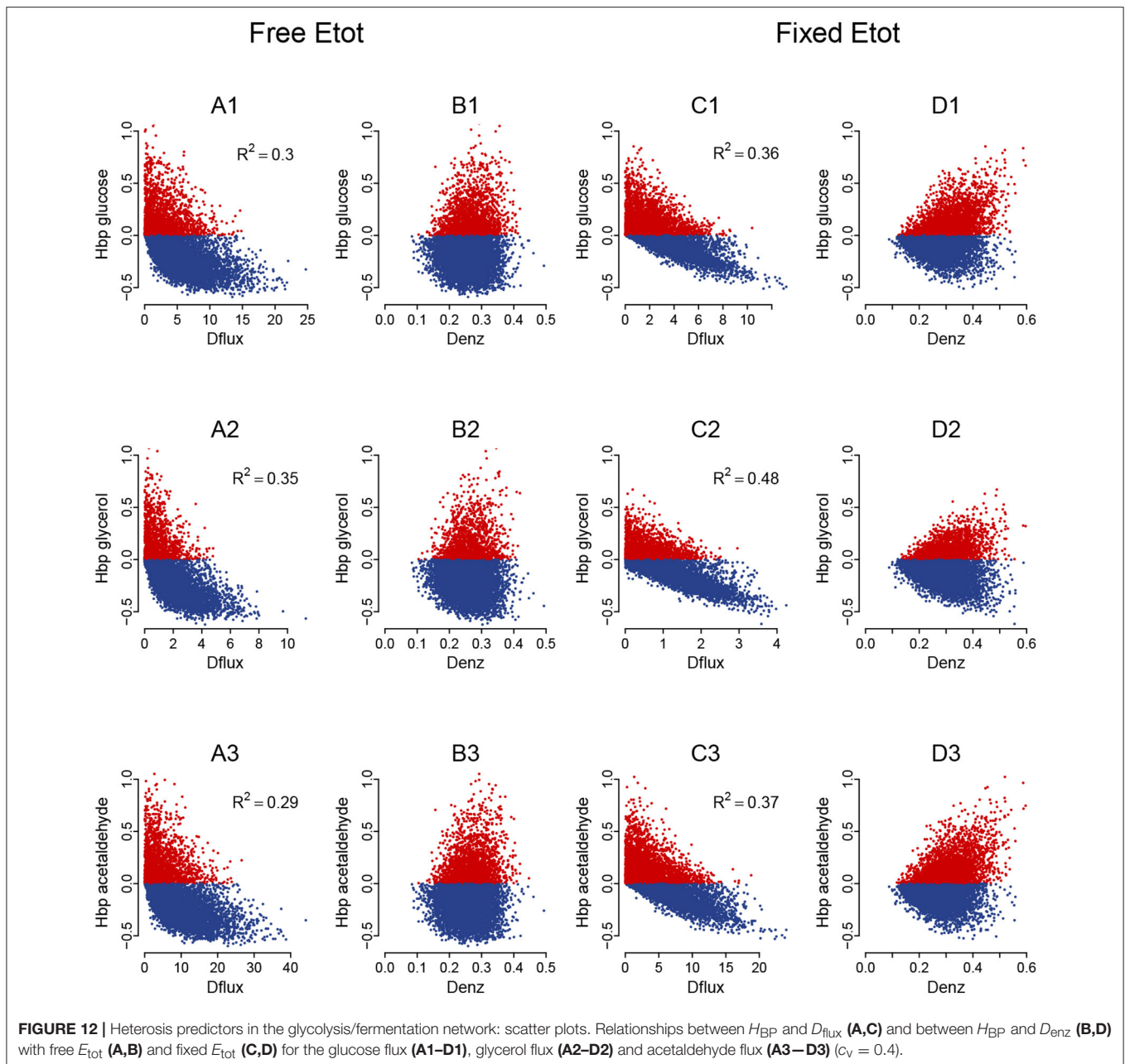


over more than one century of genetics. The bases of overdominance are diverse: (i) Pleiotropy. The individual traits controlled by the gene harbor only dominance, but the direction of dominance is reversed between the traits, which results in increased viability or fitness of heterozygotes compared to the parental homozygotes (Allison, 1954; Vrebalov et al., 2002; Lippman and Zamir, 2007); (ii) Dosage effects. When there is an optimal expression level of the gene controlling the trait, the hybrid will be heterotic whenever its expression level will be closer to the optimum than that of its parents (Hall and Wills, 1987; Flintham et al., 1997; Krieger et al., 2010); (iii) Formation of favorable oligomers in the hybrids (Tahiri-Alaoui et al., 2006; Singh et al., 2013).

Understanding heterosis when several—possibly a wealth of—polymorphic genes control the trait is challenging. However, this has been done when a limited number of genes are involved. As early as 1910, an example of heterosis accounted for by complementary dominance at two loci *L* (controlling internode length) and *T* (controlling stem thickness) was found in garden pea (Keeble and Pellew, 1910), in line with Davenport's model (Davenport, 1908), which stated that heterosis was due to recessive unfavorable alleles being masked at loci in repulsion. Since then, similar examples have been described (e.g., Warner et al., 1969). It is worth noting that many alleged overdominant QTLs (Quantitative Trait Loci) could actually correspond to pseudo-overdominance, where two (or more) linked dominant QTLs are in repulsion (Eshed and Zamir, 1995; Graham et al., 1997; Lariepe et al., 2012; Martí-Raga et al., 2017). Epistasis also accounts for heterosis through favorable intergenic interactions created in the hybrids (Powers, 1944). The textbook case (Sinnott et al., 1950), which involves both dominance and epistasis, was recognized very early on (Bateson and Gregory, 1905) and was illustrated in various plants: crossing two individuals with colorless flowers may produce hybrids with colored flowers because the parents each have a mutation inactivating a particular step of the pigment biosynthesis pathway, which disrupts the flux, and the flux is restored in the hybrid because both steps are active (e.g., Dooner et al., 1991 for anthocyanins in maize). Other cases of heterosis that have been elucidated at the molecular level also combine more than one genetic effect. For instance

in yeast, heterosis for growth at high temperature involved both dominance and epistasis at three genes *MKT1*, *END3*, and *RHO2* (Steinmetz et al., 2002b; Sinha et al., 2006), and heterosis for fermentation performance proved to be partially explained by dominance and pseudo-overdominance at four genes *VMA13*, *PDR1*, *PMA1*, and *MSB2* (Martí-Raga et al., 2017). Similar complex effects were found in the circadian clock genes *CCA1* and *LHY*, which play a central role in heterosis for chlorophyll and starch content in *A. thaliana* (Ni et al., 2009). In tomato, epistasis and overdominance at the *SFT* gene account for yield heterosis (Krieger et al., 2010).

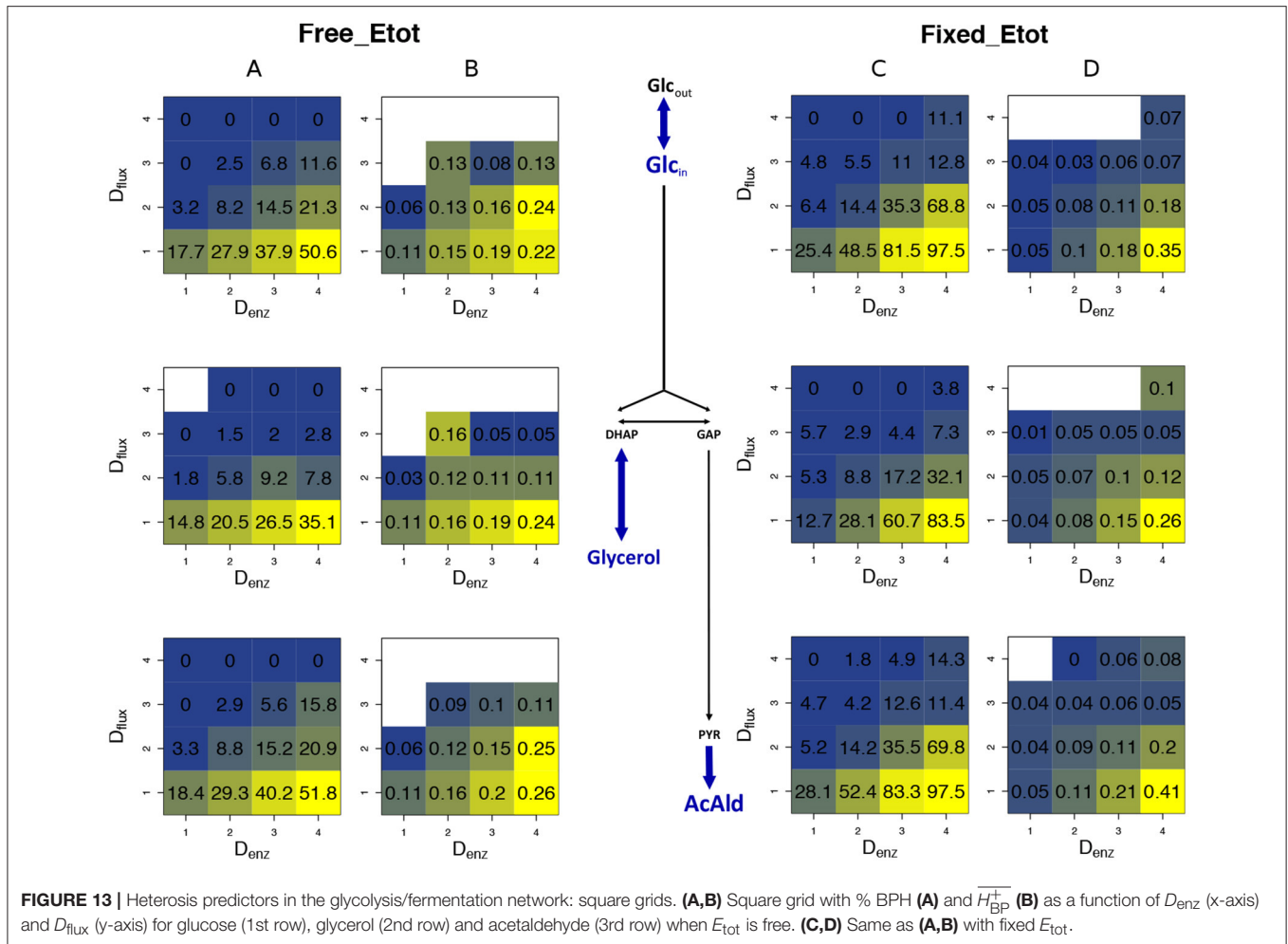
Third, given the pervasiveness of non-additive genetic effects (Veitia, 2006; Phillips, 2008), it is the *lack* of heterosis that would be a mystery. Whenever there is dominance, and/or overdominance and/or epistasis at loci governing a quantitative trait, heterosis can arise. Any one of these three genetic effects is necessary, but also sufficient, to generate heterosis. Besides, their respective importance in producing heterosis is variable, and depends on the mating system, the traits under study and the genetic material used, as shown from QTL/association mapping analyses and other approaches. In maize, dominance seems to be the main factor producing heterosis for grain yield and its components, with a modest role of overdominance and possibly epistasis (Garcia et al., 2008; Guo et al., 2014; Mezouk and Ross-Ibarra, 2014). In rice, dominance was first invoked as the main factor for heterosis of grain yield and its various components (Xiao et al., 1995), but subsequently overdominance and additive x additive or dominance x dominance epistasis have been repeatedly found for various yield-related traits (Yu et al., 1997; Li et al., 2001; Luo et al., 2001; Hua et al., 2003; Mei et al., 2005; Garcia et al., 2008; Huang et al., 2016). In Arabidopsis the main factor was claimed to be epistasis for seven growth-related traits (Melchinger et al., 2007), however, for biomass the three genetic effects seemed to be involved (Meyer et al., 2010), and a recent study showed the role of dominance and overdominance for flowering date, rosette diameter and rosette biomass (Seymour et al., 2016). In upland cotton, the three effects contribute to heterosis for yield and yield components (Shang et al., 2016). In yeast, the three effects also play a role in heterosis for growth rate measured in five different culture



media (Shapira et al., 2014), while overdominance and synergistic epistasis seem mainly involved in growth at high temperature (Shapira and David, 2016). It is worth noting that in most cases it is not possible to distinguish true overdominance from pseudo-overdominance, even if indirect arguments suggest true overdominance (Semel et al., 2006).

Despite the multiple cases where the bases of heterosis have been clarified, some authors have been searching for a “unifying principle” (Birchler et al., 2003), such as genome-wide changes in DNA methylation (Tsaftaris and Polidoros, 1999; Shen et al., 2012), small RNA expression and epigenetic regulation (Ha et al., 2009; Groszmann et al., 2011; Chen, 2013), reduced metabolic cost of protein recycling in hybrids (Goff, 2011), gene dosage effects in macro-molecular complexes

(Veitia and Vaiman, 2011), enhanced metabolic efficiency due to weak co-aggregation of allozymes (Ginn, 2017), mitochondrial complementation (McDaniel and Sarkissian, 1966; Srivastava, 1981) and phytohormonal expression (Rood et al., 1988). These molecular processes may indeed distinguish heterozygotes from homozygotes, but if they were the general hidden causes of heterosis, correlations between levels of heterosis for different traits would be observed, and this has never been reported so far (Flint-Garcia et al., 2009; Kaeppler, 2012). In fact, searching for the mechanistic bases of heterosis of a certain trait in a certain cross is equivalent to searching for the bases of its underlying genetic effects, and dominance, epistasis and overdominance have never been considered “mysterious” phenomena.



Curvature of the Genotype-Phenotype Relationship and Heterosis

The “unifying principle” proposed in this paper is not mechanistic. It is a general framework built on the common observation of the concavity of the GP relationship. The two genetic consequences of this non-linearity — dominance and epistasis — have been known for a long time in the context of the metabolic control theory (Wright, 1934; Kacser and Burns, 1981). More recently, modeling gene and metabolic networks have led to the idea that heterosis could also be a systemic property emerging from non-linear processes in the cell (Omholt et al., 2000; Fiévet et al., 2010). We provide strong arguments in support of this proposition by taking advantage of theoretical developments, and using both experimental approaches and computer simulations.

Experimentally, we performed *in vitro* 61 crosses between “parents” that differ in the concentrations of four enzymes from the upstream part of the glycolysis pathway, and observed inheritance that was biased toward positive heterosis despite additivity of enzyme concentrations in the hybrid tubes: about 44% of hybrids displayed positive MPH or BPH, whereas only 13% displayed negative MPH and none WPH. In addition, the H_{BP} index could increase up to ≈ 0.37 , which means

that with enzyme concentrations half-way between the parents, the hybrid had a flux that was 37% higher than the best parental flux. From a geometrical standpoint, these results indicate that the GP hypersurface in this *in vitro* system is globally concave, accounting for the predominance of positive heterosis, the cases of negative heterosis revealing local unevenness of the surface. We did not explore further the possible causes — metabolic and/or technical — of these irregularities.

Computer simulations based on the previous system and using a hyperbolic GP relationship allowed us to perform a large number of crosses. Enzyme concentrations were drawn, varying means and coefficients of variation, and total enzyme concentrations were either free or fixed. In all conditions only +MPH and BPH were observed, which was expected due to the concavity of the GP relationship. The results of the simulations together with our theoretical developments allowed us to identify the factors favoring heterosis:

- A contrast between parental enzyme concentrations. A high positive H_{BP} could only be observed when D_{enz} was large. Accordingly high c_v 's resulted in the highest % BPH and

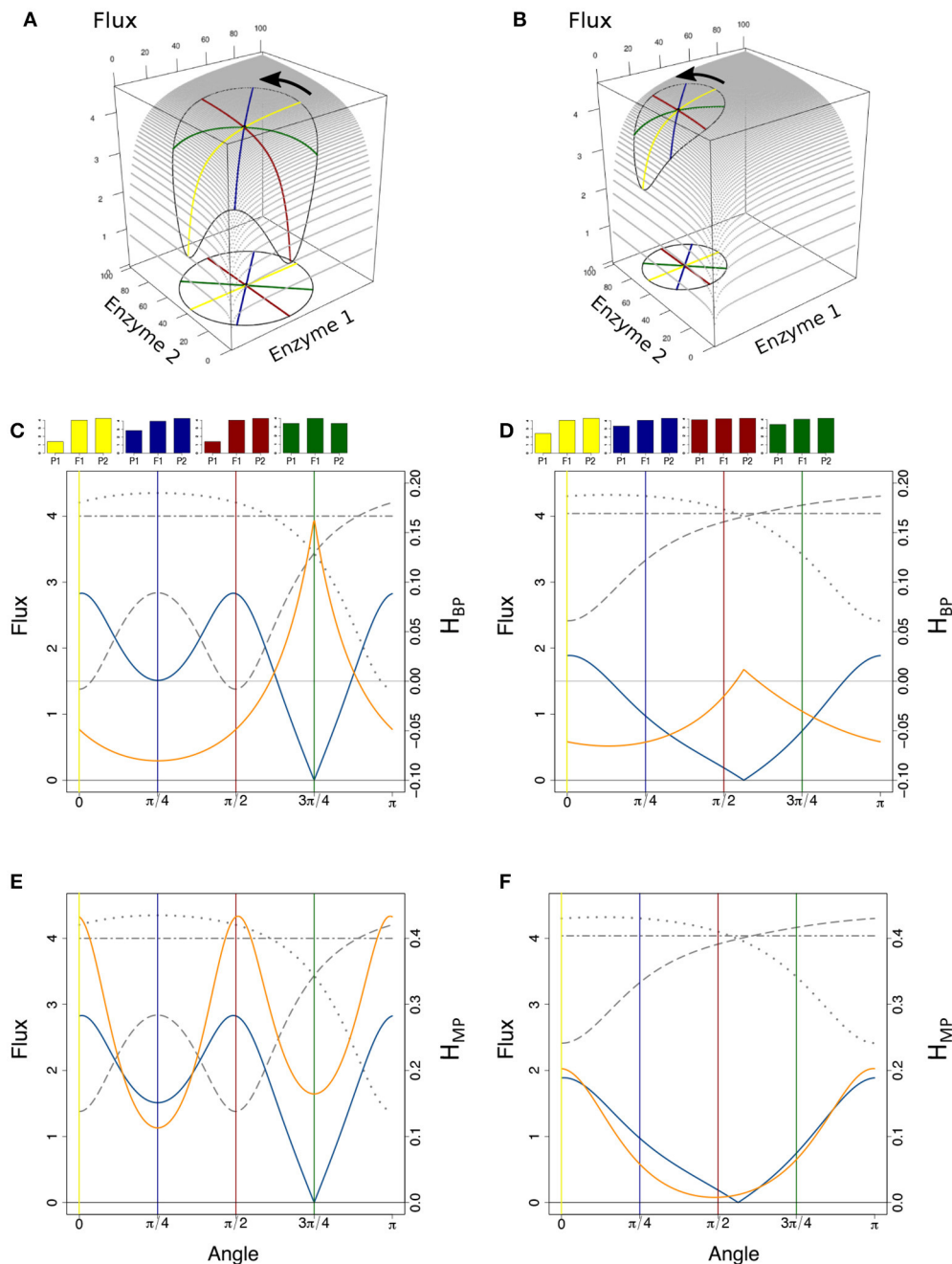


FIGURE 14 | The geometry of heterosis when the enzymatic distance between parents is constant. **(A)** Parental enzyme concentrations are on a circle of diameter $D_{enz} = 76$ centered on the hybrid point, the coordinates of which are $x = y = 40$. Four pairs of parents are highlighted by yellow, blue, red and green diameters. The closed curve on the surface shows the flux values corresponding to the circle of concentrations. The equation of the flux surface is $J = 1/(\frac{1}{E_1} + \frac{1}{E_2} + 0.2)$, in arbitrary units. **(B)** Same as **(A)**, except that $D_{enz} = 50$ and hybrid coordinates are $x = 30$ and $y = 70$. **(C)** Variation of parental fluxes J_1 and J_2 (dotted and long-dashed gray curves), of D_{flux} (blue curve) and of H_{BP} (orange curve) over the rotation of a segment joining the parents from $\alpha = 0$ (parallel to the x-axis, yellow case) to $\alpha = \pi$ (reversed positions of the parents). The two-dashed horizontal gray line corresponds to the constant hybrid flux value. The yellow, blue, red, and green vertical lines correspond to the pairs of parents of the 3D vignette in **(A)**, with barplots showing corresponding heterosis. **(D)** Same representation as in **(C)**, for the case shown in **(B)**. **(E,F)**: Same as in **(C,D)**, with H_{MP} instead of H_{BP} .

H_{BP}^+ . However, high D_{enz} is necessary but not sufficient to get BPH: the position of the parents in concentration space is also a key factor (Figure 14 and Figure S7). If a parent

contains most of the high alleles and the other most of the low alleles, BPH is unlikely because the hybrid can hardly exceed the best parent, while if the parents are complementary

for the high and low alleles, the hybrid is expected to display BPH.

- Phenotypic proximity between parents. In the case of strict concavity, BPH is very likely when the parents have close phenotypic values, and the % BPH decreases as the parental fluxes diverge. If a parental value is close to the maximum, MPH alone will be observed (**Figure 14**).
- Constraint on total enzyme concentration allocated to the system. All other things being equal, the % BPH is higher when E_{tot} is fixed, due to increased concavity of the flux response surface (**Figure 4B**).

Combining fixed E_{tot} , a large enzymatic distance and a small phenotypic difference makes BPH almost inevitable, particularly when mean enzyme concentrations are centered on their optimal values. This is consistent in geometric terms: when these conditions are met, the hybrids are closer to the “top of the hill” than their parents (**Figure S8**).

We performed similar simulations from a branched network based on glycolysis and fermentation in yeast, modeled with a system of differential equations. Three fluxes were considered: glucose input and glycerol and acetaldehyde outputs. Again BPH could only arise when the parents were phenotypically close to each other, but due to some convex enzyme-flux relations in this system, negative heterosis was also observed. For glucose and acetaldehyde fluxes, positive MPH and BPH prevailed to a large extent, since the percentage of positive heterosis varied from 69.5 to 96.9%, and for all fluxes and conditions the % BPH increased with the c_v , up to $\approx 44\%$ when E_{tot} was fixed. For glycerol, positive heterosis was predominant only for medium and high c_v 's. In all cases WPH remained very rare. Thus, the three fluxes, even though they are pleiotropically related, do not have the same type of inheritance. However, the factors favoring BPH are the same as before for the three fluxes, namely constraint on E_{tot} and a large enzymatic distance together with a small phenotypic divergence. Therefore, these factors do not depend on a particular flux-enzyme function, but only on the shape of the GP relationship. In addition we see that heterosis is a systemic property which enables a better exploitation of cell resources, since with a distribution of enzyme concentrations closer to the optimum distribution than both parents, a hybrid can “do more with less,” i.e., exceeds the best parent with a lower total enzyme concentration. In epistemological terms, it is an *emergent* property, in the sense that the individual properties of the molecular components (transcription/translation factors, mRNAs, proteins/enzymes, metabolites, etc.) cannot alone account for the phenomenon.

In order to avoid confounding effects, in particular in the glycolysis/fermentation system, we did not consider in this study possible non-additive inheritance of enzyme concentrations. Non-additivity of genetic variables has logical consequences: if it is positive, there is a higher occurrence of positive heterosis, while negative non-additivity can result in negative heterosis even in the case of a strictly concave GP relationship (Fiévet et al., 2010; Blein-Nicolas et al., 2015). In actual fact, transcripts and proteins display mostly additive inheritance (Stupar et al.,

2008), and in the case of non-additivity it is usually biased toward positive values (Swanson-Wagner et al., 2006; Hedgecock et al., 2007; Shen et al., 2012; Chen, 2013; Blein-Nicolas et al., 2015; Xing et al., 2016; Guo et al., 2017), which will reinforce positive heterosis. For certain gene/protein categories that are found to be non-additive (histone modifications, small RNAs, protein and carbon metabolism, ribosome proteins, photosynthesis, stress response and energy production), a causal link with heterosis has been suggested or evidenced (Chen, 2013; Xing et al., 2016). Incidentally, positive non-additivity at the molecular level also reflects concave GP relationships: transcript, protein and metabolite abundances are polygenic traits (Damerval et al., 1994; Brem et al., 2002; Kliebenstein, 2009) that behave like any quantitative trait and can display non-additive inheritance.

How can our results be applied to whole-cell metabolism? Crowding (Ellis, 2001) and energy cost (Dekel and Alon, 2005; Vilaprinyo et al., 2010) create constraints on the total content of cell components. For instance in yeast, Albertin et al. (2013) showed that the enzymatic pool allocated to the glycolysis, glycerol, acetate, and ethanol pathways was invariant regardless of the culture medium and strain considered. When some proteins are overexpressed, this constancy can sometimes lead to burden effects, which causes a decline in flux (Snoep et al., 1995), as we found under constrained E_{tot} . However, due to the huge number of molecular species in a cell, the increase of particular components does not usually impair the whole system. Similarly, situations of convexity within cell networks are expected to be buffered at the scale of the whole flux of matter and energy through the system. Thus, despite global constraints and local convexity, plateaued GP relationships, either S-shaped or more often concave, are commonly observed at integrated levels of cell organization, up to fitness, and most likely account for the pervasiveness of positive heterosis. Therefore, heterosis could almost be viewed as the measurement of the curvature of the GP relationship.

Our results may have practical consequences for heterosis prediction, a major issue in plant and animal breeding. A classical predictor of heterosis is the genetic distance between the parents, but the quality of the prediction depends markedly on the traits, on the genetic history of the material and on the range of genetic distances considered (Moll et al., 1965; Melchinger et al., 1992; Charcosset and Essioux, 1994; Flint-Garcia et al., 2009; Seymour et al., 2016). Our results show that better predictions of heterosis could be achieved by combining parental information at two (or more) phenotypic levels, namely the level of the trait of interest and the level (or levels) of underlying components. Omics methods allow thousands of transcripts, proteins and metabolites to be identified and quantified. Thus, we now have access to a plethora of putative predictors (Riedelsheimer et al., 2012) to be incorporated in non-linear statistical models. The high-dimensionality is challenging, but relevant variables could be selected by using both biological knowledge of the trait under study (e.g., Miller et al., 2015) and recently developed regularization methods (Dalalyan and Tsybakov, 2012; Heinzl and Tutz, 2014).

AUTHOR CONTRIBUTIONS

JF, CD, and DdV conceptualization; DdV formal analysis; CD and DdV funding acquisition; JF, TN, DdV investigation; JF, TN, CD, and DdV methodology; CD and DdV project administration; JF, TN, and DdV resources; TN and DdV software; DdV supervision, validation, visualization, writing—original draft preparation; JF, TN, CD, and DdV writing—review and editing.

ACKNOWLEDGMENTS

We dedicate this work to Pr. André Gallais, who brought us his impressive knowledge of heterosis. We thank Dr. Catherine Damerval, Dr. Michel Zivy, Dr. Mélisande Blein-Nicolas, Dr. Warren Albertin, Dr. Philippe Marullo and Dr.

François Vasseur for exciting discussions on heterosis and for useful suggestions for improving the manuscript. We also thank Marianyela Petrizzelli for her help in checking the mathematical developments and Dr. Hélène Citerne for English corrections.

This work was supported by the ANR program ALIA HeterosYeast ANR-08-ALIA-9. JF was supported by a PhD grant from the French Ministère de la Jeunesse, de l'Éducation Nationale et de la Recherche.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2018.00159/full#supplementary-material>

REFERENCES

- Agrawal, A. F., and Whitlock, M. C. (2011). Inferences about the distribution of dominance drawn from yeast gene knockout data. *Genetics* 187, 553–U274. doi: 10.1534/genetics.110.124560
- Albertin, W., Marullo, P., Bely, M., Aigle, M., Bourgaïs, A., Langella, O., et al. (2013). Linking post-translational modifications and variation of phenotypic traits. *Mol. Cell. Proteom.* 12, 720–735. doi: 10.1074/mcp.M112.024349
- Allison, A. (1954). Protection afforded by sickle-cell trait against subtertian malarial infection. *Br. Med. J.* 1, 290–294. doi: 10.1136/bmj.1.4857.290
- Barkai, N., and Leibler, S. (1997). Robustness in simple biochemical networks. *Nature* 387, 913–917. doi: 10.1038/43199
- Barth, S., Busimi, A. K., Utz, H. F., and Melchinger, A. E. (2003). Heterosis for biomass yield and related traits in five hybrids of *Arabidopsis thaliana* L. Heynh. *Heredity* 91, 36–42. doi: 10.1038/sj.hdy.6800276
- Bateson, W., and Gregory, R. P. (1905). On the inheritance of heterostylism in *Primula*. *Proc. R. Soc. Lond. Ser. B Contain. Papers Biol. Char.* 76, 581–586. doi: 10.1098/rspb.1905.0049
- Birchler, J. A., Auger, D. L., and Riddle, N. C. (2003). In search of the molecular basis of heterosis. *Plant Cell* 15, 2236–2239. doi: 10.1105/tpc.151030
- Blein-Nicolas, M., Albertin, W., Da Silva, T., Valot, B., Balliau, T., Masneuf-Pomarede, I., et al. (2015). A systems approach to elucidate heterosis of protein abundances in yeast. *Mol. Cell. Proteom.* 14, 2056–2071. doi: 10.1074/mcp.M115.048058
- Bost, B., de Vienne, D., Hospital, F., Moreau, L., and Dillmann, C. (2001). Genetic and nongenetic bases for the L-shaped distribution of quantitative trait loci effects. *Genetics* 157, 1773–1787.
- Bost, B., Dillmann, C., and de Vienne, D. (1999). Fluxes and metabolic pools as model traits for quantitative genetics. I. The L-shaped distribution of gene effects. *Genetics* 153, 2001–2012.
- Bost, B., and Veitia, R. A. (2014). Dominance and interloci interactions in transcriptional activation cascades: models explaining compensatory mutations and inheritance patterns. *Bioessays* 36, 84–92. doi: 10.1002/bies.201300109
- Brem, R. B., Yvert, G., Clinton, R., and Kruglyak, L. (2002). Genetic dissection of transcriptional regulation in budding yeast. *Science* 296, 752–755. doi: 10.1126/science.1069516
- Carey, M. (1998). The enhanceosome and transcriptional synergy. *Cell* 92, 5–8.
- Charcosset, A., and Essioux, L. (1994). The effect of population structure on the relationship between heterosis and heterozygosity at marker loci. *Theoret. Appl. Genet.* 89, 336–343.
- Charlesworth, D., and Willis, J. H. (2009). The genetics of inbreeding depression. *Nat. Rev. Genet.* 10, 783–796. doi: 10.1038/nrg2664
- Chen, Z. J. (2013). Genomic and epigenetic insights into the molecular bases of heterosis. *Nat. Rev. Genet.* 14, 471–482. doi: 10.1038/nrg3503
- Clark, A. (1991). Mutation-selection balance and metabolic control-theory. *Genetics* 129, 909–923.
- Conant, G. C., and Wolfe, K. H. (2007). Increased glycolytic flux as an outcome of whole-genome duplication in yeast. *Mol. Syst. Biol.* 3:129. doi: 10.1038/msb4100170
- Coyne, D. P. (1965). Component interaction in relation to heterosis for plant height in *Phaseolus vulgaris* L. variety crosses. *Crop Sci.* 5, 17–18.
- Dalalyan, A. S., and Tsybakov, A. B. (2012). Sparse regression learning by aggregation and Langevin Monte-Carlo. *J. Comput. Syst. Sci.* 78, 1423–1443. doi: 10.1016/j.jcss.2011.12.023
- Damerval, C., Maurice, A., Josse, J., and Devienne, D. (1994). Quantitative trait loci underlying gene-product variation - a novel perspective for analyzing regulation of genome expression. *Genetics* 137, 289–301.
- Darwin, C. (1876). *The Effects of Cross and Self Fertilisation in the Vegetable Kingdom*. London: J. Murray.
- Davenport, C. B. (1908). Degeneration, albinism and inbreeding. *Science* 28, 454–455.
- de Vienne, D., Bost, B., Fievet, J., Zivy, M., and Dillmann, C. (2001). Genetic variability of proteome expression and metabolic control. *Plant Physiol. Biochem.* 39, 271–283. doi: 10.1016/S0981-9428(01)01246-3
- Dekel, E., and Alon, U. (2005). Optimality and evolutionary tuning of the expression level of a protein. *Nature* 436, 588–592. doi: 10.1038/nature03842
- Deutschbauer, A. M., Jaramillo, D. F., Proctor, M., Kumm, J., Hillenmeyer, M. E., Davis, R. W., et al. (2005). Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics* 169, 1915–1925. doi: 10.1534/genetics.104.036871
- Dickerson, G. E. (1973). Inbreeding and heterosis in animals. *J. Anim. Sci.* 1973(Symposium), 54–77. doi: 10.1093/ansci/1973.Symposium.54
- Dooner, H. K., Robbins, T. P., and Jorgensen, R. A. (1991). Genetic and developmental control of anthocyanin biosynthesis. *Annu. Rev. Genet.* 25, 173–199. doi: 10.1146/annurev.ge.25.120191.001133
- Duvick, D. N. (2001). Biotechnology in the 1930s: the development of hybrid maize. *Nat. Rev. Genet.* 2, 69–74. doi: 10.1038/35047587
- Dykhuizen, D., Dean, A., and Hartl, D. (1987). Metabolic flux and fitness. *Genetics* 115, 25–31.
- East, E. M. (1936). Heterosis. *Genetics* 21, 375–397.
- Eldar, A., Shilo, B. Z., and Barkai, N. (2004). Elucidating mechanisms underlying robustness of morphogen gradients. *Curr. Opin. Genet. Dev.* 14, 435–439. doi: 10.1016/j.gde.2004.06.009
- Ellis, R. J. (2001). Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.* 26, 597–604. doi: 10.1016/S0968-0004(01)01938-7
- Emerson, S., Gowen, J., et al. (1952). “Biochemical models of heterosis in neurospora,” in *Heterosis* ed J. W. Gowen (Ames: Iowa State College Press), 199–217.
- Eshed, Y., and Zamir, D. (1995). An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* 141, 1147–1162.

- Felix, M.-A., and Barkoulas, M. (2015). Pervasive robustness in biological systems. *Nat. Rev. Genet.* 16, 483–496. doi: 10.1038/nrg3949
- Fell, D. (2007). *Understanding the Control of Metabolism*. London: Portland Press.
- Fiévet, J., Dillmann, C., Lagniel, G., Davantage, M., Negroni, L., Labarre, J., et al. (2004). Assessing factors for reliable quantitative proteomics based on two-dimensional gel electrophoresis. *Proteomics* 4, 1939–1949. doi: 10.1002/pmic.200300731
- Fiévet, J. B., Dillmann, C., Curien, G., and De Vienne, D. (2006). Simplified modelling of metabolic pathways for flux prediction and optimization: lessons from an *in vitro* reconstruction of the upper part of glycolysis. *Biochem. J.* 396, 317–326. doi: 10.1042/BJ20051520
- Fiévet, J. B., Dillmann, C., and de Vienne, D. (2010). Systemic properties of metabolic networks lead to an epistasis-based model for heterosis. *Theor. Appl. Genet.* 120, 463–473. doi: 10.1007/s00122-009-1203-2
- Firczuk, H., Kannambath, S., Pahle, J., Claydon, A., Beynon, R., Duncan, J., et al. (2013). An *in vivo* control map for the eukaryotic mRNA translation machinery. *Mol. Syst. Biol.* 9:635. doi: 10.1038/msb.2012.73
- Fisher, R. A. (1928). The possible modification of the response of the wild type to recurrent mutations. *Am. Natural.* 62, 115–126. doi: 10.1086/280193
- Flint-Garcia, S. A., Buckler, E. S., Tiffin, P., Ersoz, E., and Springer, N. M. (2009). Heterosis is prevalent for multiple traits in diverse maize germplasm. *PLoS ONE* 4:e7433. doi: 10.1371/journal.pone.0007433
- Flintham, J. E., Angus, W. J., and Gale, M. D. (1997). Heterosis, overdominance for grain yield, and alpha-amylase activity in F-1 hybrids between near-isogenic Rht dwarf and tall wheats. *J. Agric. Sci.* 129, 371–378. doi: 10.1017/S0021859697004899
- Garcia, A. A. F., Wang, S., Melchinger, A. E., and Zeng, Z.-B. (2008). Quantitative trait loci mapping and the genetic basis of heterosis in maize and rice. *Genetics* 180, 1707–1724. doi: 10.1534/genetics.107.082867
- Gilchrist, M. A., and Nijhout, H. F. (2001). Nonlinear developmental processes as sources of dominance. *Genetics* 159, 423–432.
- Ginn, B. R. (2017). The thermodynamics of protein aggregation reactions may underpin the enhanced metabolic efficiency associated with heterosis, some balancing selection, and the evolution of ploidy levels. *Prog. Biophys. Mol. Biol.* 126, 1–21. doi: 10.1016/j.pbiomolbio.2017.01.005
- Giorgetti, L., Siggers, T., Tiana, G., Caprara, G., Notarbartolo, S., Corona, T., et al. (2010). Noncooperative interactions between transcription factors and clustered DNA binding sites enable graded transcriptional responses to environmental inputs. *Mol. Cell* 37, 418–428. doi: 10.1016/j.molcel.2010.01.016
- Goff, S. A. (2011). A unifying theory for general multigenic heterosis: energy efficiency, protein metabolism, and implications for molecular breeding. *New Phytol.* 189, 923–937. doi: 10.1111/j.1469-8137.2010.03574.x
- Gout, J.-F., Duret, L., and Kahn, D. (2009). Differential retention of metabolic genes following whole-genome duplication. *Mol. Biol. Evol.* 26, 1067–1072. doi: 10.1093/molbev/msp026
- Gowen, J. W. (1952). *Heterosis*. Iowa State pr edition. Ames: Iowa State College Press.
- Grafius, J. (1961). Complex trait as a geometric construct. *Heredity* 16:225. doi: 10.1038/hdy.1961.24
- Grafius, J. E. (1959). Heterosis in barley. *Agron. J.* 51, 551–554. doi: 10.2134/agronj1959.00021962005100090013x
- Graham, G. I., Wolff, D. W., and Stuber, C. W. (1997). Characterization of a yield quantitative trait locus on chromosome five of maize by fine mapping. *Crop Sci.* 37, 1601–1610. doi: 10.2135/cropsci1997.0011183X003700050033x
- Groszmann, M., Greaves, I. K., Albertyn, Z. I., Scofield, G. N., Peacock, W. J., and Dennis, E. S. (2011). Changes in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to hybrid vigor. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2617–2622. doi: 10.1073/pnas.1019217108
- Guo, H., Mendrikahy, J. N., Xie, L., Deng, J., Lu, Z., Wu, J., et al. (2017). Transcriptome analysis of neo-tetraploid rice reveals specific differential gene expressions associated with fertility and heterosis. *Sci. Rep.* 7:40139. doi: 10.1038/srep40139
- Guo, T., Yang, N., Tong, H., Pan, Q., Yang, X., Tang, J., et al. (2014). Genetic basis of grain yield heterosis in an “immortalized F2” maize population. *Theor. Appl. Genet.* 127, 2149–2158. doi: 10.1007/s00122-014-2368-x
- Ha, M., Lu, J., Tian, L., Ramachandran, V., Kasschau, K. D., Chapman, E. J., et al. (2009). Small RNAs serve as a genetic buffer against genomic shock in Arabidopsis interspecific hybrids and allopolyploids. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17835–17840. doi: 10.1073/pnas.0907003106
- Hall, J., and Wills, C. (1987). Conditional overdominance at an alcohol-dehydrogenase locus in yeast. *Genetics* 117, 421–427.
- Hartl, D., Dykhuizen, D., and Dean, A. (1985). Limits of adaptation - the evolution of selective neutrality. *Genetics* 111, 655–674.
- Hedgecock, D., Lin, J.-Z., DeCola, S., Haudenschield, C. D., Meyer, E., Manahan, D. T., et al. (2007). Transcriptomic analysis of growth heterosis in larval Pacific oysters (*Crassostrea gigas*). *Proc. Natl. Acad. Sci. U.S.A.* 104, 2313–2318. doi: 10.1073/pnas.0610880104
- Heinrich, R., and Schuster, S. (1996). *The Regulation of Cellular Systems*. Boston, MA: Springer US.
- Heinrich, R., and Schuster, S. (1998). The modelling of metabolic systems. Structure, control and optimality. *Biosystems* 47, 61–77.
- Heinzel, F., and Tutz, G. (2014). Clustering in linear-mixed models with a group fused lasso penalty. *Biomet. J.* 56, 44–68. doi: 10.1002/bimj.201200111
- Hindmarsh, A. C. (1983). ODEPACK, A Systematized Collection of ODE Solvers, RS Stepleman et al.(eds.), North-Holland, Amsterdam,(vol. 1 of), pp. 55-64. *IMACS Trans. Sci. Comput.* 1, 55–64.
- Ho, W.-C., and Zhang, J. (2016). Adaptive genetic robustness of *Escherichia coli* metabolic fluxes. *Mol. Biol. Evol.* 33, 1164–1176. doi: 10.1093/molbev/msw002
- Hotz, H., Semlitsch, R. D., Gutmann, E., Guex, G.-D., and Beerli, P. (1999). Spontaneous heterosis in larval life-history traits of hemiclinal frog hybrids. *Proc. Natl. Acad. Sci. U.S.A.* 96, 2171–2176. doi: 10.1073/pnas.96.5.2171
- Houle, D. (1989). Allozyme-associated heterosis in *Drosophila melanogaster*. *Genetics* 123, 789–801.
- Hua, J., Xing, Y., Wu, W., Xu, C., Sun, X., Yu, S., et al. (2003). Single-locus heterotic effects and dominance by dominance interactions can adequately explain the genetic basis of heterosis in an elite rice hybrid. *Proc. Natl. Acad. Sci. U.S.A.* 100, 2574–2579. doi: 10.1073/pnas.0437907100
- Huang, X., Yang, S., Gong, J., Zhao, Q., Feng, Q., Zhan, Q., et al. (2016). Genomic architecture of heterosis for yield traits in rice. *Nature* 537, 629–633. doi: 10.1038/nature19760
- Hull, F. (1946). Overdominance and corn breeding where hybrid seed is not feasible. *J. Am. Soc. Agron.* 38, 1100–1103. doi: 10.2134/agronj1946.00021962003800120007x
- Izard, J., Balderas, C. D. C. G., Ropers, D., Lacour, S., Song, X., Yang, Y., et al. (2015). A synthetic growth switch based on controlled expression of RNA polymerase. *Mol. Syst. Biol.* 11:840. doi: 10.15252/msb.20156382
- Jones, D. F. (1918). The effect of inbreeding and crossbreeding upon development. *Proc. Natl. Acad. Sci. U.S.A.* 4, 246–250. doi: 10.5962/bhl.title.51006
- Kacser, H., and Burns, J. (1981). The molecular-basis of dominance. *Genetics* 97, 639–666.
- Kaeppler, S. (2011). Heterosis: one boat at a time, or a rising tide? *New Phytol.* 189, 900–902. doi: 10.1111/j.1469-8137.2010.03630.x
- Kaeppler, S. (2012). Heterosis: many genes, many mechanisms—end the search for an undiscovered unifying theory. *ISRN Bot.* 2012:682824. doi: 10.5402/2012/682824
- Kauffman, S. A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. New York, NY: Oxford University Press.
- Kazemian, M., Pham, H., Wolfe, S. A., Brodsky, M. H., and Sinha, S. (2013). Widespread evidence of cooperative DNA binding by transcription factors in *Drosophila* development. *Nucleic Acids Res.* 41, 8237–8252. doi: 10.1093/nar/gkt598
- Keeble, F., and Pellew, C. (1910). The mode of inheritance of stature and of time of flowering in peas (*Pisum sativum*). *J. Genet.* 1, 47–56. doi: 10.1007/BF02981568
- Keightley, P. D. (1996). Metabolic models of selection response. *J. Theor. Biol.* 182, 311–316. doi: 10.1006/jtbi.1996.0169
- Kliebenstein, D. (2009). Advancing genetic theory and application by metabolic quantitative trait loci analysis. *Plant Cell* 21, 1637–1646. doi: 10.1105/tpc.109.067611
- Krieger, U., Lippman, Z. B., and Zamir, D. (2010). The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nat. Genet.* 42, 459–463. doi: 10.1038/ng.550
- Lariepe, A., Mangin, B., Jasson, S., Combes, V., Dumas, F., Jamin, P., et al. (2012). The genetic basis of heterosis: multiparental quantitative trait loci mapping reveals contrasted levels of apparent overdominance among traits

- of agronomical interest in maize (*Zea mays* L.). *Genetics* 190, 795–U835. doi: 10.1534/genetics.111.133447
- Launey, S., and Hedgecock, D. (2001). High genetic load in the Pacific oyster *Crassostrea gigas*. *Genetics* 159, 255–265.
- Lewontin, R. C. (1974). *The Genetic Basis of Evolutionary Change*. New York, NY: Columbia University Press.
- Li, Z.-K., Luo, L. J., Mei, H. W., Wang, D. L., Shu, Q. Y., Tabien, R., 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.
- Lindsley, D., Sandler, L., Jacobs, P., Nozawa, H., Parry, D., Baker, B., et al. (1972). Segmental aneuploidy and genetic gross structure of drosophila genome. *Genetics* 71:157.
- Lion, S., Gabriel, F., Bost, B., Fievet, J., Dillmann, C., and de Vienne, D. (2004). An extension to the metabolic control theory taking into account correlations between enzyme concentrations. *Eur. J. Biochem.* 271, 4375–4391. doi: 10.1111/j.1432-1033.2004.04375.x
- Lippman, Z. B., and Zamir, D. (2007). Heterosis: revisiting the magic. *Trends Genet.* 23, 60–66. doi: 10.1016/j.tig.2006.12.006
- Luo, L. J., Li, Z.-K., Mei, H. W., Shu, Q. Y., Tabien, R., Zhong, D. B., et al. (2001). Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components. *Genetics* 158, 1755–1771.
- MacLean, R. C. (2010). Predicting epistasis: an experimental test of metabolic control theory with bacterial transcription and translation. *J. Evol. Biol.* 23, 488–493. doi: 10.1111/j.1420-9101.2009.01888.x
- Malmstroem, J., Beck, M., Schmidt, A., Lange, V., Deutsch, E. W., and Aebersold, R. (2009). Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans*. *Nature* 460, 762–U112. doi: 10.1038/nature08184
- Manna, F., Martin, G., and Lenormand, T. (2011). Fitness landscapes: an alternative theory for the dominance of mutation. *Genetics* 189, 923–U303. doi: 10.1534/genetics.111.132944
- Martí-Raga, M., Peltier, E., Mas, A. A., Beltran, G., and Marullo, P. (2017). Genetic causes of phenotypic adaptation to the second fermentation of sparkling wines in *Saccharomyces cerevisiae*. *G3 Genes Genom. Genet.* 7, 399–412. doi: 10.1534/g3.116.037283
- Martin, O. C., and Wagner, A. (2008). Multifunctionality and robustness trade-offs in model genetic circuits. *Biophys. J.* 94, 2927–2937. doi: 10.1529/biophysj.107.114348
- Mather, K. (1949). *Biometrical Genetics*. London: Methuen.
- McDaniel, R. G., and Sarkissian, I. V. (1966). Heterosis: Complementation by Mitochondria. *Science* 152, 1640–1642. doi: 10.1126/science.152.3729.1640
- Mei, H. W., Li, Z. K., Shu, Q. Y., Guo, L. B., Wang, Y. P., Yu, X. Q., et al. (2005). Gene actions of QTLs affecting several agronomic traits resolved in a recombinant inbred rice population and two backcross populations. *Theor. Appl. Genet.* 110, 649–659. doi: 10.1007/s00122-004-1890-7
- Melchinger, A. E., Boppenmaier, J., Dhillon, B. S., Pollmer, W. G., and Herrmann, R. G. (1992). Genetic diversity for RFLPs in European maize inbreds : II. Relation to performance of hybrids within versus between heterotic groups for forage traits. *Theor. Appl. Genet.* 84, 672–681. doi: 10.1007/BF00224167
- Melchinger, A. E., Piepho, H.-P., Utz, H. F., Muminovic, J., Wegenast, T., Törjék, O., et al. (2007). 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* 177, 1827–1837. doi: 10.1534/genetics.107.080564
- Meyer, R. C., Kusterer, B., Lisec, J., Steinfath, M., Becher, M., Scharr, H., et al. (2010). QTL analysis of early stage heterosis for biomass in Arabidopsis. *Theor. Appl. Genet.* 120, 227–237. doi: 10.1007/s00122-009-1074-6
- Mezmouk, S., and Ross-Ibarra, J. (2014). The pattern and distribution of deleterious mutations in maize. *G3 (Bethesda)* 4, 163–171. doi: 10.1534/g3.113.008870
- Miller, M., Song, Q., Shi, X., Juenger, T. E., and Chen, Z. J. (2015). Natural variation in timing of stress-responsive gene expression predicts heterosis in intraspecific hybrids of Arabidopsis. *Nat. Comm.* 6:7453. doi: 10.1038/ncomms8453
- Moll, R. H., Lonnquist, J. H., Fortuno, J. V., and Johnson, E. C. (1965). The relationship of heterosis and genetic divergence in maize. *Genetics* 52:139.
- Moreno-Sanchez, R., Saavedra, E., Rodríguez-Enriquez, S., and Olin-Sandoval, V. (2008). Metabolic control analysis: a tool for designing strategies to manipulate metabolic pathways. *J. Biomed. Biotechnol.* 2008:597913. doi: 10.1155/2008/597913
- Ni, Z., Kim, E.-D., Ha, M., Lackey, E., Liu, J., Zhang, Y., et al. (2009). Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature* 457, 327–331. doi: 10.1038/nature07523
- Nijhout, H. F., Berg, A. M., and Gibson, W. T. (2003). A mechanistic study of evolvability using the mitogen-activated protein kinase cascade. *Evol. Dev.* 5, 281–294. doi: 10.1046/j.1525-142X.2003.03035.x
- Nijhout, H. F., Best, J. A., and Reed, M. C. (2015). Using mathematical models to understand metabolism, genes, and disease. *Bmc Biol.* 13:79. doi: 10.1186/s12915-015-0189-2
- Omholt, S. W., Plahte, E., Oyehaug, L., and Xiang, K. F. (2000). Gene regulatory networks generating the phenomena of additivity, dominance and epistasis. *Genetics* 155, 969–980.
- O'Neill, D. G., Church, D. B., McGreevy, P. D., Thomson, P. C., and Brodbelt, D. C. (2013). Longevity and mortality of owned dogs in England. *Vet. J.* 198, 638–643. doi: 10.1016/j.tvjl.2013.09.020
- Orr, H. (1991). A Test of Fisher Theory of Dominance. *Proc. Natl. Acad. Sci. U.S.A.* 88, 11413–11415.
- Perfeito, L., Ghozzi, S., Berg, J., Schnetz, K., and Laessig, M. (2011). Nonlinear fitness landscape of a molecular pathway. *PLoS Genet.* 7:e1002160. doi: 10.1371/journal.pgen.1002160
- Perry, M. W., Bothma, J. P., Luu, R. D., and Levine, M. (2012). Precision of hunchback expression in the drosophila embryo. *Curr. Biol.* 22, 2247–2252. doi: 10.1016/j.cub.2012.09.051
- Phillips, P. C. (2008). Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* 9, 855–867. doi: 10.1038/nrg2452
- Poker, G., Zarai, Y., Margaliot, M., and Tuller, T. (2014). Maximizing protein translation rate in the non-homogeneous ribosome flow model: a convex optimization approach. *J. Roy. Soc. Interface* 11:UNSP 20140713. doi: 10.1098/rsif.2014.0713
- Powers, L. (1944). An expansion of Jones's theory for the explanation of heterosis. *Am. Natural.* 78, 275–280. doi: 10.1086/281199
- Pryciak, P. M. (2008). Customized signaling circuits. *Science* 319, 1489–1490. doi: 10.1126/science.1156414
- Rauscher, M. D. (2013). The evolution of genes in branched metabolic pathways. *Evolution* 67, 34–48. doi: 10.1111/j.1558-5646.2012.01771.x
- R Core Team (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available online at: <http://www.R-project.org/>
- Rendel, J. M. (1962). The relationship between gene and phenotype. *J. Theor. Biol.* 2, 296–308. doi: 10.1016/0022-5193(62)90032-2
- Rest, J. S., Morales, C. M., Waldron, J. B., Oplente, D. A., Fisher, J., Moon, S., et al. (2013). Nonlinear fitness consequences of variation in expression level of a eukaryotic gene. *Mol. Biol. Evol.* 30, 448–456. doi: 10.1093/molbev/mss248
- Richey, F. D. (1942). Mock-dominance and hybrid vigor. *Science* 96, 280–281. doi: 10.1126/science.96.2490.280
- Riedelsheimer, C., Czedik-Eysenberg, A., Grieder, C., Lisec, J., Technow, F., Sulpice, R., et al. (2012). Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nat. Genet.* 44, 217–220. doi: 10.1038/ng.1033
- Roberts, H. F. (1965). *Plant Hybridization Before Mendel*. New York, NY: Hafner publishing co Ltd.
- Rodrigues, J. V., Bershtein, S., Li, A., Lozovsky, E. R., Hartl, D. L., and Shakhnovich, E. I. (2016). Biophysical principles predict fitness landscapes of drug resistance. *Proc. Natl. Acad. Sci. U.S.A.* 113, E1470–E1478. doi: 10.1073/pnas.1601441113
- Rohde, P., Hinch, D. K., and Heyer, A. G. (2004). Heterosis in the freezing tolerance of crosses between two *Arabidopsis thaliana* accessions (Columbia-0 and C24) that show differences in non-acclimated and acclimated freezing tolerance. *Plant J.* 38, 790–799. doi: 10.1111/j.1365-3113.2004.02080.x
- Rood, S., Buzzell, R., Mander, L., Pearce, D., and Pharis, R. (1988). Gibberellins - a Phytohormonal Basis for Heterosis in Maize. *Science* 241, 1216–1218. doi: 10.1126/science.241.4870.1216

- Rossignol, R., Faustin, B., Rocher, C., Malgat, M., Mazat, J. P., and Letellier, T. (2003). Mitochondrial threshold effects. *Biochem. J.* 370, 751–762. doi: 10.1042/bj20021594
- Schnell, F. W., and Cockerham, C. C. (1992). Multiplicative vs. arbitrary gene action in heterosis. *Genetics* 131, 461–469.
- Semel, Y., Nissenbaum, J., Menda, N., Zinder, M., Krieger, U., Issman, N., et al. (2006). Overdominant quantitative trait loci for yield and fitness in tomato. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12981–12986. doi: 10.1073/pnas.0604635103
- Seymour, D. K., Chae, E., Grimm, D. G., Martín Pizarro, C., Habring-Müller, A., Vasseur, F., et al. (2016). Genetic architecture of nonadditive inheritance in *Arabidopsis thaliana* hybrids. *Proc. Natl. Acad. Sci. U.S.A.* 113, E7317–E7326. doi: 10.1073/pnas.1615268113
- Shang, L., Wang, Y., Cai, S., Wang, X., Li, Y., Abduweli, A., et al. (2016). Partial dominance, overdominance, epistasis and QTL by environment interactions contribute to heterosis in two upland cotton hybrids. *G3 Genes Genom. Genet.* 6, 499–507. doi: 10.1534/g3.115.025809
- Shapira, R., and David, L. (2016). Genes with a combination of over-dominant and epistatic effects underlie heterosis in growth of *Saccharomyces cerevisiae* at high temperature. *Front. Genet.* 7:72. doi: 10.3389/fgene.2016.00072
- Shapira, R., Levy, T., Shaked, S., Fridman, E., and David, L. (2014). Extensive heterosis in growth of yeast hybrids is explained by a combination of genetic models. *Heredity* 113, 316–326. doi: 10.1038/hdy.2014.33
- Shen, H., He, H., Li, J., Chen, W., Wang, X., Guo, L., et al. (2012). Genome-wide analysis of DNA methylation and gene expression changes in two *Arabidopsis* ecotypes and their reciprocal hybrids. *Plant Cell* 24, 875–892. doi: 10.1105/tpc.111.094870
- Shull, G. H. (1914). Duplicate gene for capsule-form in *Bursa bursa-pastoris*. *Zeitschrift Ind. Abst. Vererbungsl.* 12, 97–149.
- Singh, R., Low, E.-T. L., Ooi, L. C.-L., Ong-Abdullah, M., Ting, N.-C., Nagappan, J., et al. (2013). The oil palm SHELL gene controls oil yield and encodes a homologue of SEEDSTICK. *Nature* 500, 340–344. doi: 10.1038/nature12356
- Sinha, H., Nicholson, B. P., Steinmetz, L. M., and McCusker, J. H. (2006). Complex genetic interactions in a quantitative trait locus. *PLoS Genet.* 2:e13. doi: 10.1371/journal.pgen.0020013
- Sinnot, E. W., Dunn, L. C., and Dobzhansky, T. (1950). *Principles of genetics*. New York, NY; Toronto; London: McGraw-Hill Book Co. Inc.
- Snoep, J. L., Yomano, L. P., Westerhoff, H. V., and Ingram, L. O. (1995). Protein burden in *Zymomonas mobilis*: negative flux and growth control due to overproduction of glycolytic enzymes. *Microbiology* 141, 2329–2337. doi: 10.1099/13500872-141-9-2329
- Sprague, G. F. (1983). “Heterosis in maize: theory and practice,” in *Heterosis, Monographs on Theoretical and Applied Genetics*, Vol 6., ed R. Frankel (Berlin; Heidelberg: Springer), 47–70. doi: 10.1007/978-3-642-81977-3_2
- Srivastava, H. K. (1981). Intergenomic interaction, heterosis, and improvement of crop yield. *Adv. Agron.* 34, 117–195. doi: 10.1016/S0065-2113(08)60886-X
- Steinmetz, L. M., Scharfe, C., Deutschbauer, A. M., Mokranjac, D., Herman, Z. S., Jones, T., et al. (2002a). Systematic screen for human disease genes in yeast. *Nat. Genet.* 31, 400–404. doi: 10.1038/ng929
- Steinmetz, L. M., Sinha, H., Richards, D. R., Spiegelman, J. I., Oefner, P. J., McCusker, J. H., et al. (2002b). Dissecting the architecture of a quantitative trait locus in yeast. *Nature* 416, 326–330. doi: 10.1038/416326a
- Stupar, R. M., Gardiner, J. M., Oldre, A. G., Haun, W. J., Chandler, V. L., and Springer, N. M. (2008). Gene expression analyses in maize inbreds and hybrids with varying levels of heterosis. *BMC Plant Biol.* 8:33. doi: 10.1186/1471-2229-8-33
- 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 F1 hybrid and its inbred parents. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6805–6810. doi: 10.1073/pnas.0510430103
- Szathmari, E. (1993). Do Deleterious mutations act synergistically - metabolic control-theory provides a partial answer. *Genetics* 133, 127–132.
- Tahiri-Alaoui, A., Sim, V. L., Caughey, B., and James, W. (2006). Molecular heterosis of prion protein β -oligomers. A potential mechanism of human resistance to disease. *J. Biol. Chem.* 281, 34171–34178. doi: 10.1074/jbc.M606606200
- Teusink, B., Passarge, J., Reijenga, C. A., Esgalhado, E., van der Weijden, C. C., Schepper, M., et al. (2000). Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *Eur. J. Biochem.* 267, 5313–5329. doi: 10.1046/j.1432-1327.2000.01527.x
- Tsaftaris, A. S., and Polidoros, A. N. (1999). DNA methylation and plant breeding. *Plant Breed. Rev.* 18, 87–176. doi: 10.1002/9780470650158.ch3
- Vasseur, F., Violle, C., Enquist, B. J., Granier, C., and Vile, D. (2012). A common genetic basis to the origin of the leaf economics spectrum and metabolic scaling allometry. *Ecol. Lett.* 15, 1149–1157. doi: 10.1111/j.1461-0248.2012.01839.x
- Veitia, R. A. (2006). *The Biology of Genetic Dominance*. Gerogetown, TX: Landes Bioscience.
- Veitia, R. A., and Vaiman, D. (2011). Exploring the mechanistic bases of heterosis from the perspective of macromolecular complexes. *FASEB J.* 25, 476–482. doi: 10.1096/fj.10-170639
- Vilaprinyo, E., Alves, R., and Sorribas, A. (2010). Minimization of biosynthetic costs in adaptive gene expression responses of yeast to environmental changes. *PLoS Comput. Biol.* 6:e1000674. doi: 10.1371/journal.pcbi.1000674
- Vrebalov, J., Ruezinsky, D., Padmanabhan, V., White, R., Medrano, D., Drake, R., et al. (2002). A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus. *Science* 296, 343–346. doi: 10.1126/science.1068181
- Wagner, A. (2013). *Robustness and Evolvability in Living Systems*. Princeton; Oxford: Princeton University Press.
- Warner, R. L., Hageman, R. H., Dudley, J. W., and Lambert, R. J. (1969). Inheritance of nitrate reductase activity in *Zea mays* L. *Proc. Natl. Acad. Sci. U.S.A.* 62, 785–792. doi: 10.1073/pnas.62.3.785
- West, G. B., Brown, J. H., and Enquist, B. J. (1999). The fourth dimension of life: Fractal geometry and allometric scaling of organisms. *Science* 284, 1677–1679. doi: 10.1126/science.284.5420.1677
- Whitacre, J. M. (2012). Biological robustness: paradigms, mechanisms, and systems principles. *Front. Genet.* 3:67. doi: 10.3389/fgene.2012.00067
- Wright, S. (1934). Physiological and evolutionary theories of dominance. *Am. Natural.* 68, 24–53. doi: 10.1086/280521
- Xiao, J., Li, J., Yuan, L., and Tanksley, S. (1995). Dominance is the major genetic-basis of heterosis in rice as revealed by Qtl analysis using molecular markers. *Genetics* 140, 745–754.
- Xing, J., Sun, Q., and Ni, Z. (2016). Proteomic patterns associated with heterosis. *BBA-Prot. Proteom.* 1864, 908–915. doi: 10.1016/j.bbapap.2015.12.007
- Yang, A., Emmerson, D. A., Dunnington, E. A., and Siegel, P. B. (1999). Heterosis and developmental stability of body and organ weights at hatch for parental line broiler breeders and specific crosses among them. *Poultry Sci.* 78, 942–948. doi: 10.1093/ps/78.7.942
- Yu, S. B., Li, J. X., Xu, C. G., Tan, Y. F., Gao, Y. J., Li, X. H., et al. (1997). Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc. Natl. Acad. Sci. U.S.A.* 94, 9226–9231. doi: 10.1073/pnas.94.17.9226
- Zarai, Y., Margaliot, M., and Tuller, T. (2014). Maximizing protein translation rate in the ribosome flow model: the homogeneous case. *IEEE-Acm Trans. Comput. Biol. Bioinform.* 11, 1184–1195. doi: 10.1109/TCBB.2014.2330621
- Zorgo, E., Gjuvsland, A., Cubillos, F. A., Louis, E. J., Liti, G., Blomberg, A., et al. (2012). Life history shapes trait heridity by accumulation of loss-of-function alleles in yeast. *Mol. Biol. Evol.* 29, 1781–1789. doi: 10.1093/molbev/mss019

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Fiévet, Nidelet, Dillmann and de Vienne. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.