


Commentary

# The Pitfalls of Heterosis Coefficients

Dominique de Vienne \*  and Julie B. Fiévet

GQE—Le Moulon, INRAE, Université Paris-Saclay, CNRS, AgroParisTech, 91190 Gif-sur-Yvette, France; julie.fievet@inrae.fr

\* Correspondence: dominique.de-vienne@inrae.fr; Tel.: +33-1-69332360

Received: 29 May 2020; Accepted: 4 July 2020; Published: 9 July 2020



**Abstract:** Heterosis (hybrid vigour) is a universal phenomenon of crucial agro-economic and evolutionary importance. We show that the most common heterosis coefficients do not properly measure deviation from additivity because they include both a component accounting for “real” heterosis and a term that is not related to heterosis, since it is derived solely from parental values. Therefore, these coefficients are inadequate whenever the aim of the study is to compare heterosis levels between different traits, environments, genetic backgrounds, or developmental stages, as these factors may affect not only the level of non-additivity, but also parental values. The only relevant coefficient for such comparisons is the so-called “potence ratio”. Because most heterosis studies consider several traits/stages/environmental conditions, our observations support the use of the potence ratio, at least in non-agronomic contexts, because it is the only non-ambiguous heterosis coefficient.

**Keywords:** flowering; growth rate; heterosis measurement; hybrid vigour; maize; non-additivity; potence ratio; yield

## 1. Introduction

Non-linear processes are extremely common in biology. In particular, genotype-phenotype or phenotype-phenotype relationships frequently display concave behaviours, resulting in the dominance of “high” over “low” alleles [1] and in positive heterosis for a wide range of polygenic traits [2,3]. Properly quantifying the degree of non-additivity is an essential prerequisite for interpreting and comparing genetic studies and for making predictions in plant and animal breeding. In this commentary paper, we first recap the different ways non-additivity is measured in genetics. Subsequently, we analyse the formal relationships between the different heterosis coefficients and provide examples drawn from experimental studies in maize and cotton. Finally, we show the extent to which the most commonly used heterosis coefficients may lead to interpretation errors.

## 2. The Dominance and Heterosis Coefficients

There are two classical coefficients to measure the degree of dominance:

(i) Wright [1] defined:

$$D_W = \frac{z_2 - z_{12}}{z_2 - z_1}$$

where  $z_1$ ,  $z_2$  and  $z_{12}$  are, respectively, the phenotypic values of genotypes  $A_1A_1$ ,  $A_2A_2$  and  $A_1A_2$ , with  $z_2 > z_1$ .  $D_W$  varies from 0, when  $A_2$  is fully dominant over  $A_1$ , to 1, when  $A_2$  is fully recessive with respect to  $A_1$ .  $D_W = 0.5$  corresponds to semi-dominance or additivity ( $z_{12} = \frac{z_1+z_2}{2}$ ) (Table 1). Note that  $D_W$  is strictly equivalent to the dominance coefficient  $h$  used in evolutionary genetics [4].



The  $D_W$  and  $D_F$  coefficients are linearly related:

$$D_F = 1 - 2D_W$$

Thus, dominance can be quantified with either coefficient, since both of them give the position of the heterozygote relative to the parental homozygotes.

For polygenic traits, either coefficient could be used to quantify non-additivity, i.e., “real” heterosis, without any ambiguity. Actually, one finds five heterosis coefficients in the literature (see their characteristic values in Table 1).

The two most popular coefficients are the best-parent (BP) and mid-parent (MP) heterosis coefficients (e.g., [6,7]):

$$H_{BP} = \frac{z_{12} - z_2}{z_2}$$

$$H_{MP} = \frac{z_{12} - \bar{z}}{\bar{z}}$$

where  $z_2$ ,  $z_{12}$  and  $\bar{z}$  are, respectively, the phenotypic values of the parent 2 (with  $z_2 > z_1$ ), of the parent 1  $\times$  parent 2 hybrid and of the parental mean.

In some instances, the authors do not normalize the difference between the hybrid and the best- or mid-parent value. Fonseca & Patterson [8] proposed:

$$H_{bp} = z_{12} - z_2$$

and Falconer [5]:

$$H_{mp} = z_{12} - \bar{z}$$

Finally, the so-called “potence ratio” [9] has the same expression as Falconer’s dominance coefficient ( $D_F$ ):

$$H_{PR} = \frac{z_{12} - \bar{z}}{\frac{z_2 - z_1}{2}}$$

A value of 0 indicates additivity, 1 indicates  $z_{12} = z_2$  (hybrid value = best-parent value),  $-1$  indicates  $z_{12} = z_1$  (hybrid value = worst-parent value), and  $> 1$  (resp.  $< -1$ ) indicates best-parent (resp. worst-parent) heterosis (Table 1).  $H_{PR}$  explicitly includes the values of the three genotypes, whereas the other coefficients lack one of the parental values ( $H_{BP}$  and  $H_{bp}$ ) or both ( $H_{MP}$  and  $H_{mp}$ —a given mean can correspond to an infinity of parental values). From a genetic point of view,  $H_{PR}$  is explicitly expressed in terms of the five genetic effects contributing to heterosis (Supplementary Table S1). Thus, the potence ratio, which is still by far the least used heterosis coefficient, is the only one that informs us of the exact position of the hybrid value relative to the parental values. Wright’s dominance coefficient has the same property, but its inverse direction of variation, which makes comparisons less easy, probably explains why it is not used in this context.

### 3. Relationship between the Potence Ratio and the other Heterosis Coefficients

It is easy to show that the relationship between  $H_{PR}$  and the other coefficients is (with  $z_2 > z_1$ ):

$$H_{MP} = H_{PR}z_m \quad (1)$$

$$H_{BP} = \left( \frac{-1 + H_{PR}}{2} \right) z_b \quad (2)$$

$$H_{mp} = H_{PR}\bar{z}z_m \quad (3)$$

$$H_{bp} = \left( \frac{-1 + H_{PR}}{2} \right) z_2 z_b \quad (4)$$

where  $z_m = \frac{z_2 - z_1}{z_1 + z_2}$  (the difference between the parental values normalized by their sum) and  $z_b = \frac{z_2 - z_1}{z_2}$  (the difference between the parental values normalized by the best parental value).

For a given  $H_{PR}$  value, the coefficients  $H_{MP}$  and  $H_{BP}$  are linearly related to  $z_m$  and  $z_b$ , respectively, i.e., they depend on the scale of the parental values. More specifically, the relationship between  $H_{MP}$  and  $z_m$  is negative when  $H_{PR} < 0$  and positive when  $H_{PR} > 0$ , while the relationship between  $H_{BP}$  and  $z_b$  is negative when  $H_{PR} < 1$  and positive when  $H_{PR} > 1$ . As  $z_m$  and  $z_b$  are positive, we see from Equation (1) that for  $H_{PR} \neq 0$ , we have

$$0 < H_{MP} < H_{PR} \quad \text{if } H_{PR} > 0 \quad (\text{positive mid-parent heterosis})$$

and

$$H_{PR} < H_{MP} < 0 \quad \text{if } H_{PR} < 0 \quad (\text{negative mid-parent heterosis})$$

Regarding  $H_{BP}$ , we see from Equation (2) that, for  $H_{PR} \neq 1$ , we have

$$0 < H_{BP} < \left( \frac{-1 + H_{PR}}{2} \right) \quad \text{if } H_{PR} > 1 \quad (\text{positive best-parent heterosis})$$

and

$$\left( \frac{-1 + H_{PR}}{2} \right) < H_{BP} < 0 \quad \text{if } H_{PR} < 1 \quad (\text{negative best-parent heterosis})$$

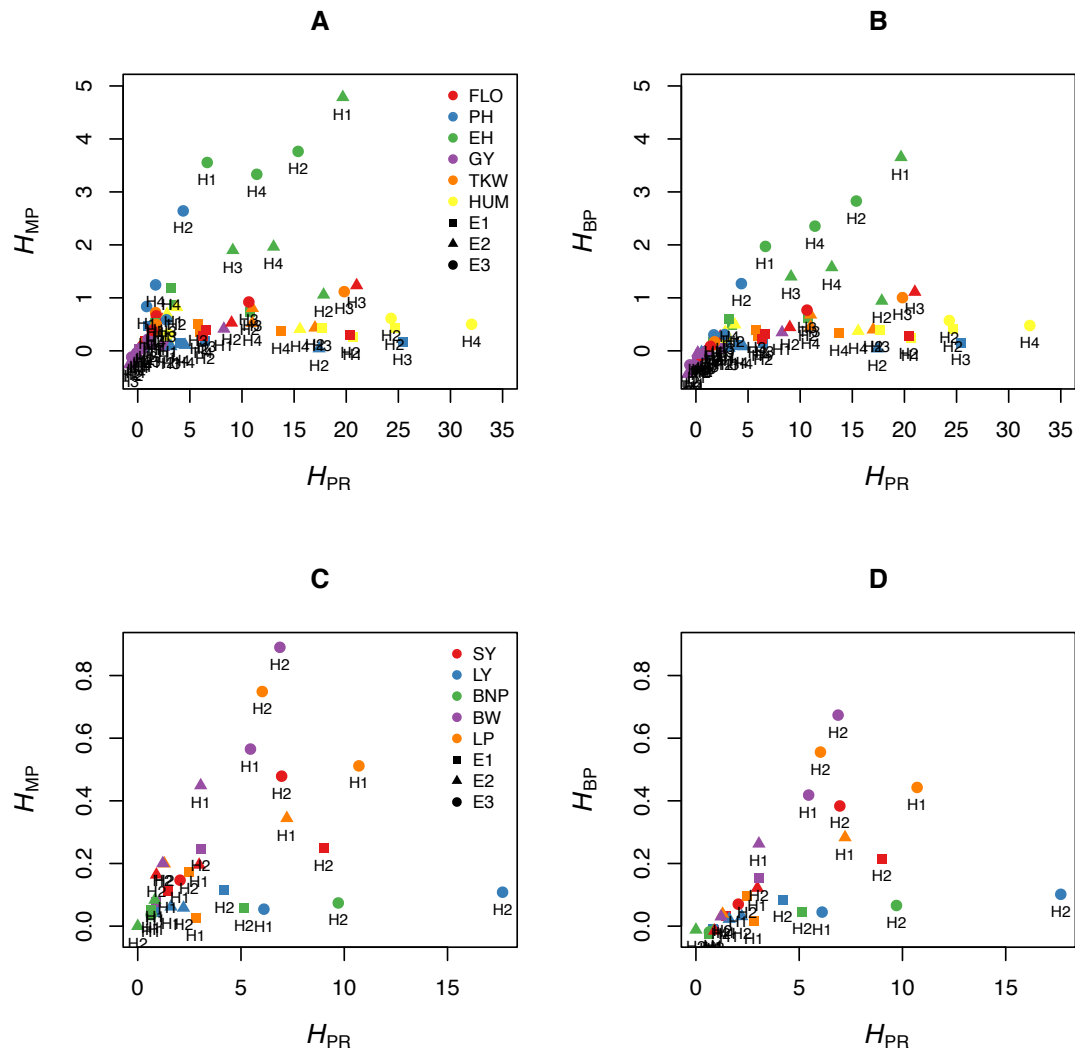
If  $H_{PR} = 0$  (resp.  $H_{PR} = 1$ ),  $H_{MP}$  (resp.  $H_{BP}$ ) = 0.

Numerical applications performed with nine  $H_{PR}$  values, from  $H_{PR} = -2$  to  $H_{PR} = 2$ , show that the same  $H_{MP}$  or  $H_{BP}$  value can be observed for very different  $H_{PR}$  values, depending on  $z_m$  and  $z_b$  values, respectively (Supplementary Figure S1). For instance,  $H_{MP} \approx 0.4$  can correspond to both mid-parent heterosis ( $H_{PR} = 0.5$ ,  $z_m \approx 0.8$ ) and best-parent heterosis ( $H_{PR} = 2$ ,  $z_m \approx 0.21$ ).

We illustrate this by using experimental data from maize. We measured six traits (flowering time, plant height, ear height, grain yield, thousand-kernel weight, and kernel moisture) in four crosses (B73  $\times$  F252, F2  $\times$  EP1, F252  $\times$  EP1, F2  $\times$  F252) grown in three different environments in France (Saint-Martin-de-Hinx in 2014, Jargeau in 2015, and Rhodon in 2015). We computed  $H_{PR}$ ,  $H_{MP}$  and  $H_{BP}$  for the 72 trait-cross-environment combinations. Figure 1A,B shows that the relationship between  $H_{PR}$  and the other two coefficients is very loose, if any. A given  $H_{PR}$  value can correspond to a wide range of  $H_{MP}$  or  $H_{BP}$  values, and vice versa. We performed the same analyses using the data published by Shang et al. [10], who measured five traits in two crosses of cotton grown in three environments. The same loose relationship between  $H_{PR}$  and either heterosis coefficient was observed (Figure 1C,D). This means that the normalized differences between the parents, which are not related to heterosis, since they do not include values from the hybrids, markedly affect  $H_{MP}$  and  $H_{BP}$ .

Regarding the  $H_{mp}$  and  $H_{bp}$  coefficients, which are not dimensionless, they only provide the direction of heterosis. For a given  $H_{PR}$  value,  $H_{mp}$  can vary from  $-\infty$  to 0 when  $H_{PR} < 0$  and from 0 to  $+\infty$  when  $H_{PR} > 0$ , and  $H_{bp}$  can vary from  $-\infty$  to 0 when  $H_{PR} < 1$  and from 0 to  $+\infty$  when  $H_{PR} > 1$  (Equations (3) and (4)).

Let us examine the possible interpretation errors that may result from the use of  $H_{MP}$  and  $H_{BP}$ .



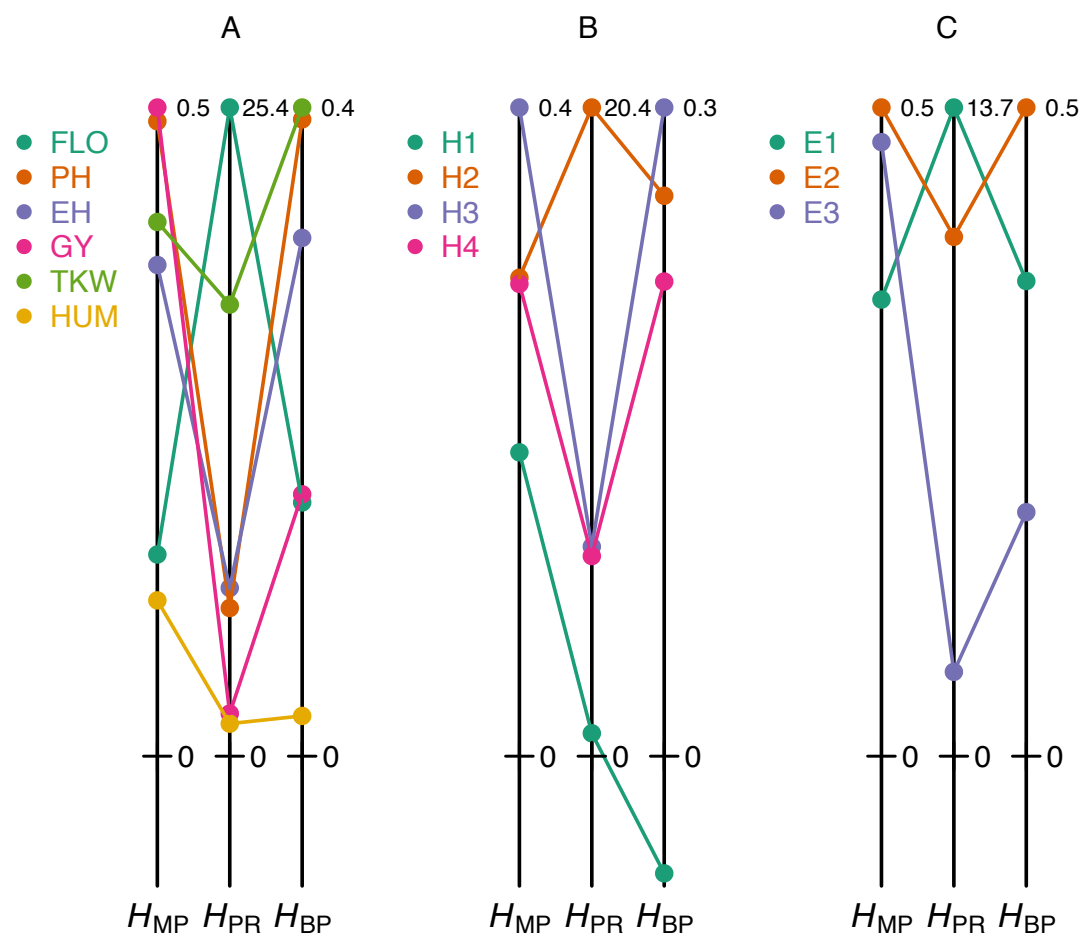
**Figure 1.** Relationship between the potence ratio  $H_{PR}$  and the two heterosis coefficients  $H_{MP}$  and  $H_{BP}$ . (A,B) Six traits were measured in maize (FLO: flowering time [days between 50% flowering and 12 August], PH: plant height, EH: ear height, GY: grain yield, TKW: thousand-kernel weight, KM: kernel moisture) in four crosses (H1: B73  $\times$  F252, H2: F2  $\times$  EP1, H3: F252  $\times$  EP1, H4: F2  $\times$  F252) grown in three environments in France (E1: Saint-Martin-de-Hinx, E2: Jargeau, E3: Rhodon). (A) Relationship between  $H_{PR}$  and  $H_{MP}$ . (B) Relationship between  $H_{PR}$  and  $H_{BP}$ . (For clarity, four of the 72 trait-cross-environment combinations are not represented because they have high  $H_{PR}$  values.) (C,D) Five traits were measured in cotton (SY: seed yield [grams per plant], LY: lint yield [grams per plant], BNP: bolls per plant, BW: boll weight [grams], LP: lint percentage) in two crosses (H1: X1135  $\times$  GX100-2 and H2: GX1135  $\times$  VGX100-2) grown in three environments in China (E1: Handan, E2: Cangzhou, E3: Xiangyang) (Data from [10]). (C) Relationship between  $H_{PR}$  and  $H_{MP}$ . (D) Relationship between  $H_{PR}$  and  $H_{BP}$ .

#### 4. The Pitfalls of the Most Commonly Used Heterosis Coefficients

The non-univocal relationship between  $H_{PR}$  and the most commonly used heterosis coefficients has two consequences. (i) Comparing the coefficient values for a given trait in different crosses and/or environments and/or developmental stages leads to erroneous conclusions whenever these factors have an effect on the scale of the trait and/or on the difference between the parental values (i.e., on  $z_m$  or  $z_b$ ). Possible differences in deviations from additivity between these conditions cannot be detected. (ii) This problem is even more pronounced when studying different traits, because each trait has

its own scale of variation, making  $H_{MP}$  and  $H_{BP}$  (and to a greater extent  $H_{mp}$  and  $H_{bp}$ ) useless for comparing the real levels of heterosis of these traits.

These pitfalls can easily be illustrated from our maize dataset. Figure 2A shows that classifying a set of traits according to their degree of heterosis can give markedly different results, depending on whether one uses the  $H_{PR}$  coefficient or one of the two coefficients  $H_{MP}$  and  $H_{BP}$ . For instance, in the F252  $\times$  EP1 cross, flowering time displays moderate heterosis according to  $H_{MP}$  and  $H_{BP}$  even though this trait actually has the highest  $H_{PR}$  value. Conversely, plant height is the second most heterotic trait according to  $H_{MP}$  and  $H_{BP}$ , but not according to  $H_{PR}$ . Similarly, comparing heterosis of a given trait in different hybrids results in coefficient-specific rankings: heterosis of ear height measured with  $H_{PR}$  is highest in the B73  $\times$  F252 hybrid, whereas according to  $H_{MP}$  and  $H_{BP}$  the highest values are found in the F252  $\times$  EP1 hybrid (Figure 2B). Finally, the effect of the environment on heterosis also reveals obvious discrepancies between  $H_{PR}$  on the one hand and  $H_{MP}$  or  $H_{BP}$  on the other hand (Figure 2C).

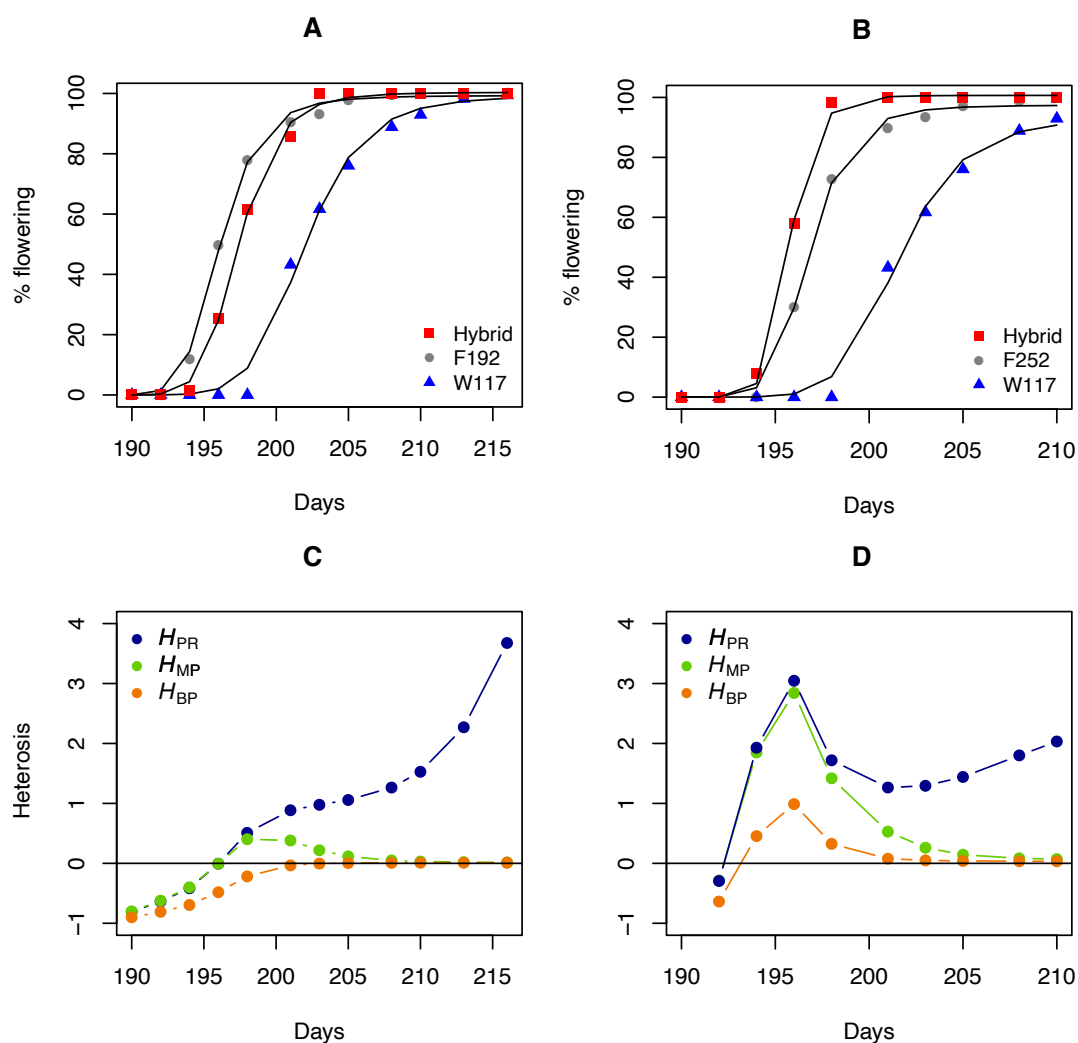


**Figure 2.** Heterosis values obtained with different coefficients. (A) Heterosis coefficients for six traits measured in the F252  $\times$  EP1 cross grown in Saint-Martin-de-Hinx (France) in 2014. (B) Heterosis coefficients for ear height in four crosses grown in Saint-Martin-de-Hinx (France) in 2014. (C) Heterosis coefficients for plant height in the F2  $\times$  F252 cross grown in the three environments. The six traits and the three environments are the same as in Figure 1A. The scales of the heterosis coefficients are normalized by the maximum value in each dataset (figures at the top right of the vertical lines).

It is also informative to compare the profiles of heterosis coefficients for a trait measured during development or growth. A Hill function was used to fit the percentage of flowering individuals over time in the W117 × F192 and W117 × F252 hybrids and their parents:

$$y = \frac{ax^n}{b + x^n}$$

where  $x$  is the time,  $a$  and  $b$  are constants, and  $n$  is the Hill coefficient. We then computed the heterosis coefficients over time for the percentage of flowering individuals estimated from the fitted curves (Figure 3). Again,  $H_{PR}$  tells a different story when compared to  $H_{MP}$  and  $H_{BP}$ . Because hybrid and parental values converge as flowering progresses, both  $H_{MP}$  and  $H_{BP}$  inevitably decrease when flowering nears 100%. The evolution of  $H_{PR}$ , which quantifies the "real" heterosis, is clearly different, with a monotonous increase in the W117 × F192 hybrid and a fluctuating profile in the W117 × F252 hybrid.



**Figure 3.** Heterosis for flowering in two maize hybrids. (A) Percentage of flowering over time (number of days since 1 January) for parents W117 and F192 and their hybrid, adjusted with a Hill function. (B) Percentage of flowering over time for parents W117 and F252 and their hybrid. (C) Heterosis coefficient profiles over time for the W117 × F192 cross. (D) Heterosis coefficient profiles over time for the W117 × F252 cross.

Similar results were observed in a simulation describing the increase in population size of a unicellular organism, which exhibits logistic growth. We used:

$$y = \frac{K}{1 + ae^{-r\theta}}$$

where  $y$  is the size of the population,  $K$  the carrying capacity,  $a$  a constant,  $r$  the growth rate, and  $\theta$  the time. We assumed that the parents only differed in their growth rate  $r$  and that there was additivity for this parameter. The results show that the  $H_{MP}$  and  $H_{BP}$  profiles for population size over time are clearly not congruent with that of  $H_{PR}$  (Supplementary Figure S2).

## 5. Discussion

Using simple theoretical considerations and relying on experimental data and simulations, we showed that the most commonly used heterosis coefficients, i.e.,  $H_{MP}$  and  $H_{BP}$  (and their non-normalized forms  $H_{mp}$  and  $H_{bp}$ ) cannot and should not be used if the heterosis levels are to be compared between different traits, environments, genetic backgrounds, or developmental stages. Because their expression does not explicitly include the two parental values in addition to the hybrid value, these coefficients, unlike the potence ratio  $H_{PR}$ , do not quantify the deviation from additivity but only the normalized distance between the hybrid value and either the best or the mean parental value. The extent to which erroneous conclusions can be drawn when performing comparisons using these coefficients was illustrated with data from maize and cotton, and from population growth simulation in a micro-organism.

If  $H_{MP}$ ,  $H_{BP}$ ,  $H_{mp}$ , and  $H_{bp}$  do not provide reliable information on non-additivity, why are they so commonly used? There are probably both historical and technical reasons: (i) the first scientists who quantified heterosis were plant breeders [11,12]. From an economic perspective, the goal was, and still is, to develop hybrids that are “better” than the best- or mid-parent values for the desired agronomic traits, and not to know where the hybrid value is relative to the parental values. Heterosis coefficients have been defined accordingly and the habit has remained; (ii) the coefficients giving the right non-additivity values,  $H_{PR}$  for heterosis and  $D_W$  or  $D_F$  for dominance, can take on high to very high values when the parents are close, due to the small differences  $z_2 - z_1$  in the denominator of the fractions. This can produce extreme values that are not easy to represent and manipulate for statistical treatments. Nevertheless, such values are biological realities that precisely convey the inheritance of the traits under study, something that  $H_{MP}$ ,  $H_{BP}$ ,  $H_{mp}$ , and  $H_{bp}$  do not. Note that the two dominance coefficients used for monogenic traits have the same property, which does not prevent their use to the exclusion of any other. In addition, from a practical point of view, a single coefficient is sufficient to know the position of the hybrid relative to the mid- or best-parent, whereas in a number of studies the authors compute and comment both  $H_{MP}$  and  $H_{BP}$  (or  $H_{mp}$  and  $H_{bp}$ ). In conclusion, to compare the amplitude of heterosis between traits, developmental stages, crosses, or environmental conditions, there is no other choice but to use the only heterosis coefficient— $H_{PR}$ —that is not affected by the scale of the parental values and that accounts for the position of the hybrid in the parental range.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2223-7747/9/7/875/s1>, Table S1. Heterosis coefficients expressed as functions of genetic effects. Figure S1. Influence of the scale of the parental values on  $H_{MP}$  and  $H_{BP}$  for different values of the potence ratio  $H_{PR}$ . Figure S2. Heterosis for population size (simulations).

**Author Contributions:** Conceptualization: D.d.V. Maize experiments: J.B.F. Data analyses and numerical applications: D.d.V. and J.B.F. D.d.V. and J.B.F. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** The experiments in maize were funded by the French Agence Nationale de la Recherche (Amazing project ANR-10-BTBR-01).



**Acknowledgments:** We thank our colleagues Mélisande Blein-Nicolas, Michel Zivy, Judith Legrand, Christine Dillmann and Charlotte Coton for their helpful comments on the manuscript, and H el ene Citerne for English corrections. We are grateful to the key persons at the INRA Station of Saint-Martin-de-Hinx, of Euralis and of MASseeds for the 2014 and 2015 field experiments.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Wright, S. Physiological and evolutionary theories of dominance. *Am. Nat.* **1934**, *68*, 24–53. [[CrossRef](#)]
2. Fi evet, J.B.; Nidelet, T.; Dillmann, C.; de Vienne, D. Heterosis is a systemic property emerging from non-linear genotype-phenotype relationships: Evidence from in vitro genetics and computer simulations. *Front. Genet.* **2018**, *9*, 159. [[CrossRef](#)] [[PubMed](#)]
3. Vasseur, F.; Fouqueau, L.; de Vienne, D.; Nidelet, T.; Violle, C.; Weigel, D. Nonlinear phenotypic variation uncovers the emergence of heterosis in *Arabidopsis thaliana*. *PLoS Biol.* **2019**, *17*, e3000214. [[CrossRef](#)] [[PubMed](#)]
4. Crow, J.F.; Kimura, M. *An Introduction to Population Genetics Theory*; The Blackburn Press: Caldwell, NJ, USA, 1970.
5. Falconer, D.S. *Introduction to Quantitative Genetics*; Oliver and Boyd: Edinburgh, UK; London, UK, 1960.
6. Gowen, J.W. *Heterosis*; Iowa State College Press: Ames, IA, USA, 1952.
7. Frankel, R. *Heterosis: Reappraisal of Theory and Practice*; Springer: Berlin/Heidelberg, Germany; New York, NY, USA; Tokyo, Japan, 1983.
8. Fonseca, S.; Patterson, F.L. Hybrid Vigor in a Seven-Parent Diallel Cross in Common Winter Wheat (*Triticum aestivum* L.). *Crop. Sci.* **1968**, *8*, 85–88. [[CrossRef](#)]
9. Mather, K. *Biometrical Genetics*; Methuen: London, UK, 1949.
10. Shang, L.; Wang, Y.; Cai, S.; Wang, X.; Li, Y.; Abduweli, A.; Hua, J. Partial Dominance, Overdominance, Epistasis and QTL by Environment Interactions Contribute to Heterosis in Two Upland Cotton Hybrids. *G3 Genes Genomes Genet.* **2016**, *6*, 499–507. [[CrossRef](#)] [[PubMed](#)]
11. Shull, G.H. The composition of a field of maize. *Am. Breeder's Assoc. Rep.* **1908**, *4*, 296–301. [[CrossRef](#)]
12. East, E.M. Heterosis. *Genetics* **1936**, *21*, 375–397. [[PubMed](#)]



  2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).