

1 **Non-linear phenotypic variation uncovers the emergence of**
2 **heterosis in *Arabidopsis thaliana***

3 François Vasseur^{1,2,3*}, Louise Fouqueau², Dominique de Vienne⁴, Thibault Nidelet⁵, Cyrille
4 Violle² and Detlef Weigel^{1*}

5 ¹Max Planck Institute for Developmental Biology, D-72076 Tübingen, Germany

6 ²CEFE, CNRS, Univ Montpellier, Université Paul Valéry Montpellier 3, EPHE, IRD,
7 Montpellier, France

8 ³Laboratoire d'Ecophysiologie des Plantes sous Stress Environnementaux (LEPSE), INRA,
9 Montpellier SupAgro, UMR759, F-34060 Montpellier, France

10 ⁴GQE–Le Moulon, INRA, Université Paris-Sud, CNRS, AgroParisTech, Université
11 Paris-Saclay, 91190 Gif-sur-Yvette, France

12 ⁵Sciences Pour l'Énologie, INRA, Université de Montpellier, Montpellier SupAgro,
13 34060 Montpellier, France

14 *Corresponding authors. Corresponding authors' addresses:

15 franc.vasseur@gmail.com

16 weigel@weigelworld.org

17 **Short title:** Allometric relationships and non-additive inheritance

18 **Abstract**

19 Heterosis describes the phenotypic superiority of hybrids over their parents in traits related to
20 fitness. Understanding and predicting non-additive inheritance such as heterosis is crucial for
21 evolutionary biology, as well as for plant and animal breeding. However, the physiological bases
22 of heterosis remain debated. Moreover, empirical data in various species have shown that
23 diverse genetic and molecular mechanisms are likely to explain heterosis, making it difficult to
24 predict its emergence and amplitude from parental genotypes alone. In this study, we evaluated
25 a model of physiological dominance proposed by Sewall Wright to explain the non-additive
26 inheritance of metabolic fluxes at the cellular level. We used 450 hybrids derived from crosses
27 among natural inbred accessions of *Arabidopsis thaliana* to test Wright's model for two fitness-
28 related traits at the whole-plant level: growth rate and fruit number. We found that allometric
29 relationships between traits constrain phenotypic variation in hybrids and inbreds to a similar
30 extent. These allometric relationships behave predictably, in a non-linear manner, explaining up
31 to 75% of heterosis amplitude, while genetic distance among parents at best explains 7%. Thus,
32 our findings are consistent with Wright's model of physiological dominance on plant
33 performance, and suggest that the emergence of heterosis is an intrinsic property of non-linear
34 relationships between traits. Furthermore, our study highlights the potential of a geometric
35 approach of phenotypic relationships for predicting heterosis of two major components of crop
36 productivity and yield.

37 **Key Words:** Heterosis; hybrid vigour; molecular dominance; non-linear relationships; biomass;
38 metabolic scaling; non-additive inheritance; allometry; genetic distance; phenotypic distance;
39 growth rate; fruit production; yield; phenotypic integration; physiology; predictive model.

40 Introduction

41 If the inheritance of phenotypic traits was always additive, progenies will always exhibit
42 intermediate trait values compared to their respective parents. Pure genetic additivity is,
43 however, the exception rather than the rule in intraspecific crosses. Non-additive inheritance has
44 been exploited for decades in agronomy [1,2], although the underlying mechanisms remain a
45 major question for evolutionary genetics and crop science [2,3]. In plants, hybrid vigour or
46 heterosis has been frequently observed [4–9], and its molecular bases have been investigated in
47 numerous studies [9–13]. Unfortunately, empirical observations often led to contrasting
48 conclusions, and none of the proposed genetic mechanisms can fully explain the emergence of
49 heterosis of different traits across all study systems [9,14–19]. Thus, we are still lacking a
50 unifying theoretical corpus that enables us to explain and predict heterosis of fitness-related
51 traits, including biomass, growth rate and reproductive success.

52 Several genetic hypotheses have been proposed to explain heterosis [3,10,13,20].
53 According to the ‘dominance hypothesis’, each parent contributes favourable, dominant alleles
54 (generally at many loci) that together complement the deleterious effects of recessive alleles
55 originating from the other parent. The ‘overdominance hypothesis’ postulates the existence of
56 loci where the heterozygous state (Aa) is superior to both homozygotes (AA or aa). With
57 overdominance, the emergence of a hybrid superior to its best parent could depend on a single
58 locus. Pseudo-overdominance corresponds to cases where dominant, favourable alleles are in
59 linkage with recessive, unfavourable alleles, so that the heterozygous combinations appear to
60 behave as overdominant loci. The picture is further complicated by the contributions of epistasis
61 [15,21] and epigenetics [12,22] to heterosis in plants.

62 To tease apart the different hypotheses, quantitative trait locus (QTL) mapping has been
63 carried out in many species [5,9,15,17,21,23]. In general, individual studies differ strongly in
64 their conclusions regarding the underlying genetic mechanisms, apparently depending on the
65 investigated traits, the genetic material used, the mating system, and the experimental

66 constraints related to the number of crosses necessary for robust statistical inference (*e.g.*, diallel
67 mating design) [13,19,20]. In addition, the QTL approach performs poorly to detect small-effect
68 loci [24]. Yet, it has been shown that fitness-related traits such as growth rate, size and fruit
69 production are often controlled by a large number of genes, which individually may have very
70 weak effects [25–27].

71 Because of the polygenic nature of fitness-related traits, heterosis is expected to be
72 associated with molecular dominance, and should positively correlate with genetic distance
73 between parents, at least up to a certain extent [28–30]. Some findings are consistent with this
74 hypothesis [31–33], suggesting that parental genetic distance could be used to quantitatively
75 predict heterosis in plants. Unfortunately, experimental studies generally employ for practical
76 reasons only a relative small number of crosses and parental lines [7,34], which makes it difficult
77 to generalize the findings of individual studies. In recent work, Seymour and colleagues [13]
78 investigated heterosis in a half-diallel between 30 genome-sequenced accessions of *Arabidopsis*
79 *thaliana* collected from diverse Eurasian populations [35]. As expected under the dominance
80 model [28–30], they found a positive correlation between parental genetic distance and
81 heterosis. However, genetic distance between parents only accounted for less than 3% of
82 heterosis among *A. thaliana* hybrids, making predictions based on genetic distances alone
83 strongly uncertain.

84 Strikingly, many continue to consider the physiological bases of heterosis a mystery
85 [36,37]. An alternative to genetics-first studies in order to understand and predict heterosis is to
86 consider the physiological constraints that determine phenotypic variation across organisms. As
87 early as 1934, Sewall Wright proposed a model of physiological dominance to explain the
88 maintenance of recessive alleles in natural populations [38]. Wright began with the universal
89 relationship that connects the concentration of enzymes to the metabolic flux that results from
90 their activity (Fig. 1). Since the relationship between these two traits is concave with a
91 horizontal asymptote (*e.g.*, Michaelis-Menten kinetics), dominance of metabolic flux is expected
92 even if enzyme concentrations are additive and hybrids intermediate between their parents (Fig.

93 1). Modern knowledge of metabolic networks was incorporated into this model by Kacser and
94 Burns in 1981 [39], but Wright's model of dominance has otherwise received little attention in
95 the heterosis literature. Recently, Fiévet and colleagues [40,41] evaluated Wright's model with a
96 focus on the relationship between metabolic flux and enzyme activity in the chain of glycolysis
97 in yeast. They modelled the importance of the curvature of the relationship between enzyme
98 concentration and metabolic flux in the emergence of heterosis on glycolysis output [40]. This
99 approach has considerable promise for predicting the phenotype of hybrids by considering the
100 non-linear relationships that often link phenotypic traits with each other. However, these
101 theoretical expectations remain to be tested on more complex traits, as well as on multicellular
102 organisms such as plants, with potentially major perspectives for cultivated species.

103 Many phenotypic relationships exhibit non-linear geometries, notably between fitness-
104 related traits [42,43]. These relationships are analogous to those between metabolic and
105 enzymatic activities described by Wright [38], Kacser and Burns [39], and Fiévet and colleagues
106 [40,41]. In particular, comparative ecology studies have demonstrated the existence of non-
107 linear allometric relationships between the biomass of an organism and its morphology,
108 physiology and metabolism [42,44–46]. Allometric relationships often take the form of a
109 function $Y = \alpha M^\beta$, where Y is a morphological or physiological trait (*e.g.*, respiration rate, growth
110 rate, leaf area, reproductive biomass) that can be predicted by M , the biomass of the organism, a
111 constant α and the allometric exponent β [44,47]. The allometric exponent of many traits
112 exhibits a value less than one, causing a concavity of the relationship (called 'strict allometry'
113 [42]). Allometric relationships are usually linearized by logarithmic transformation, especially to
114 compare exponent values (slope of the log-log relationship) between organisms [48]. For
115 instance, the metabolic scaling theory predicts that the allometric exponent is a multiple of a
116 quarter (*e.g.*, $\frac{1}{4}$, $\frac{3}{4}$) for many biological processes such as respiration rate, hydraulic resistance,
117 carbon fluxes, biomass allocation and growth rate [44,46,49–51]. Recent studies with plants
118 [43,52] suggest that the exponent must be corrected for the organism's biomass because it is
119 linked to the adaptive diversification of plant metabolism [27].

120 In this study, we tested, inspired by Wright's model of physiological dominance for
121 enzyme concentration and metabolic flux [38], whether non-linear allometric relationships can
122 explain the emergence of heterosis in macroscopic traits related to performance, such as growth
123 rate and fruit production. We used the plant *A. thaliana*, which is not a crop, but has been used
124 for many genetic studies of heterosis [7,13,53–56]. This species is characterized by a high rate of
125 inbreeding, which leads to a high rate of homozygosity in natural populations [57,58]. The
126 complete sequencing of 1,135 natural accessions [59] has provided abundant information on the
127 extent of genetic variation between populations, the genotype-phenotype map of the species, as
128 well as the causes of phenotypic variation in intraspecific hybrids [60–64]. Furthermore,
129 allometric relationships that link growth rate and fruit production to plant biomass have been
130 evaluated in this species [27,52,65]. Here, we specifically addressed the following two questions:
131 (i) Are there non-linear allometric relationships that similarly constrain trait variation in natural
132 accessions and hybrids? And (ii) do these non-linear relationships explain the emergence and
133 extent of heterosis in growth rate and fruit production? To this end, we compared several traits
134 between 451 inbred accessions representing a wide range of native environments and 450
135 crosses derived from crosses among 415 of these accessions. Our results are consistent with
136 many aspects of Wright's model, suggesting that heterosis emerges intrinsically from the non-
137 linear relationships between traits.

138 **Results**

139 **Trait variation and heterosis in a wide range of *A. thaliana* genotypes**

140 We generated 450 hybrids by manual crosses between 415 inbred accessions (Fig. 2A). We then
141 measured four traits, vegetative dry mass, plant lifespan, growth rate and total fruit number, in
142 the 450 hybrids, the 415 inbred parents, plus a further 35 inbred accessions. The parental
143 combinations for the hybrids had been chosen to represent a wide range of genetic, geographic
144 and phenotypic distances (Fig. 2B,C).

145 Traits were strongly variable between genotypes: vegetative dry mass M varied between
146 1.3 and 2,218 mg, plant lifespan between 24 and 183 days, growth rate between 0.04 and 40.4
147 mg d^{-1} , and fruit number between 5 and 400 (Table S1). Across the entire dataset of hybrids and
148 inbreds, a high proportion of the phenotypic variance could be explained by genotypic
149 differences (broad-sense heritability): from 80% for growth rate to 94% for vegetative dry mass
150 (Table S2). We found similar ranges of variation for both inbreds and hybrids (Fig. 3). For
151 instance, standard deviation of plant lifespan was 19.8 days across inbreds, and 20.1 days across
152 hybrids (Table S2). Hybrids were on average not significantly different from inbreds ($P > 0.01$
153 for all traits, Table S2).

154 Despite the similar distribution of trait values for inbreds and hybrids, we found
155 significant heterosis for all traits. This was measured by the discrete categorization of hybrid
156 phenotypic class based on the comparison of hybrid trait distribution with both mid-parent
157 values and highest ('best') or lowest ('worst') parent values. There was both positive and
158 negative heterosis (Fig. 3). However, the extent and direction of heterosis was variable between
159 traits. For instance, 21% of hybrids had confidence intervals (CIs) for lifespan that significantly
160 differed from both or mean parental value, and thus, were considered as significantly heterotic.
161 By contrast, 56% of hybrids were heterotic for vegetative dry mass. In general, there was more
162 negative than positive heterosis, specifically for vegetative dry mass (46% of hybrids). Thus,
163 hybrids represent a similar phenotypic space as natural inbred accessions, and the emergence of
164 heterosis in our data did not simply reflect a few exceptional hybrids that outperformed all
165 inbreds.

166 **Similar non-linear allometric relationships in accessions and hybrids**

167 In this study, we focused on two traits with important evolutionary and agronomic outcomes:
168 growth rate and fruit number. As predicted by ecological theory, these two traits exhibited non-
169 linear allometric relationships with vegetative dry mass (Fig. 4). Consistent with recent studies
170 of plant allometry [27,43,52], growth rate could be modelled by a power-law function with
171 mass-corrected exponent ($g(M)$, Table 1), with a concave curvature (Fig. 4A). By contrast, fruit

172 number exhibited a right-skewed bell-shaped relationship (Fig. 4B). We modelled this allometric
173 relationship with an inverse quadratic function ($f(M)$, Table 1). Importantly, inbreds and hybrids
174 exhibited similar non-linear relationships, characterized by coefficients that were not
175 significantly different from each other (with overlapping 95% CIs; Table 1).

176 **Relationships between heterosis and genetic, phenotypic and geographic distances**

177 We first compared what fraction of heterosis could be explained by genetic and phenotypic
178 distances between inbred parents. For all hybrids, we quantified heterosis as the observed
179 phenotypic deviation relative to mid-parent value (MPH), and best parental value (BPH).
180 Pairwise genetic distances were calculated either with all SNPs in the genome, or with SNPs in
181 the 1% top-genes associated with the corresponding trait [27]. For the phenotypic distance, we
182 used the absolute difference in vegetative dry mass between parents.

183 We found that genetic distance between parents was positively correlated with heterosis
184 of growth rate (both MPH and BPH $P < 0.001$; Fig. 5A). However, genetic distance only
185 accounted for less than 10% of heterosis (7% and 6% for MPH and BPH, respectively).
186 Phenotypic distance was also poorly correlated with heterosis of growth rate, and only with
187 MPH ($r^2 = 0.03$, Fig. 5B). By contrast, genetic distance did not correlate with heterosis of fruit
188 number (both $P > 0.01$, Fig. 5C), while phenotypic distance was negatively correlated, albeit
189 poorly, with BPH ($r^2 = 0.05$, Fig. 5D). Using only the 1% top-genes associated with growth rate
190 and fruit number, which are more likely to have a causal role in the measured traits, did not
191 improve correlations ($r^2 < 0.10$). Even when combined in a multivariate model, genetic and
192 phenotypic distances together explained less than 10% of heterosis of growth rate and fruit
193 number (Table S3). For comparison, geographic distance between parents explained less than
194 2% of heterosis of fruit number, and less than 0.2% of heterosis of growth rate. In summary,
195 none of the parental distances – be they genetic, phenotypic or geographic – had much power to
196 explain and predict a significant portion of heterosis.

197 **Explaining variation in heterosis by phenotypic non-linearity**

198 In a second approach, we used the fitted equations (Table 1) to take into account the non-
199 linearity of allometric relationships and to predict heterosis. Our first goal was to predict growth
200 rate and fruit number with (i) the allometric relationship fitted on parents, and (ii) the
201 measurement of vegetative dry mass in hybrids (Fig. 6). Predicted growth rate in hybrids
202 strongly correlated with observed growth rate ($r^2 = 0.95$, Fig. S1A). By contrast, predicted fruit
203 number was poorly correlated with observed trait values ($r^2 = 0.08$, Fig. S1B). This is consistent
204 with the larger dispersion of trait values around the fitted curve for fruit number (Fig. 4B)
205 compared to growth rate (Fig. 4A). We then compared trait deviation predicted in hybrids by
206 non-linearity of traits – relative to the predicted mid-parent value (PNL_{MP}) and to the predicted
207 best-parent value (PNL_{BP}) – with the observed MPH and BPH (Fig. 6). Observed heterosis and
208 predicted non-linearity of growth rate were strongly correlated ($r^2 = 0.75$ and 0.66 for PNL_{MP} vs
209 MPH and PNL_{BP} vs BPH, respectively, Fig. 6B). Observed heterosis and predicted phenotypic
210 non-linearity of fruit number were also positively correlated, although weaker than for growth
211 rate ($r^2 = 0.14$ and 0.10 for PNL_{MP} vs MPH and PNL_{BP} vs BPH, respectively, Fig. 6D). This
212 suggests that allometric relationships allow the prediction of heterosis amplitude, and that
213 prediction accuracy depends on the strength of the underlying non-linear covariation between
214 traits.

215 **Discussion**

216 Already in 1934, Wright wrote in his seminal paper that “dominance has to do with the
217 physiology of the organism and has nothing to do with the mechanism of transmission” [38].
218 Eighty years later, the emergence of heterosis is still considered as an enigma and its
219 physiological bases remain debated. Despite the many dominant, overdominant and epistatic
220 QTL identified in a plethora of species [19,20,36], none of the genetic models has been formally
221 validated in more than a few cases [*e.g.*, 9]. Here, we approached the question of heterosis from
222 a physiological angle based on a geometric analysis of trait-trait relationships. Combining the

223 ideas put forth by Wright [38] with principles from allometric theory [44,51], we demonstrate
224 that a significant part of heterosis can be explained by the non-linear relationships that link
225 phenotypic traits with each other.

226 The accuracy of heterosis prediction with phenotypic non-linearity depends on how
227 strongly traits correlate with each other. Metabolic scaling theory [44,46,66] postulates that
228 body size constrains trait variation in the form of universal mathematical laws. According to the
229 postulates of this theory, biomass determines the number of metabolically active units (cells,
230 mitochondria), which in turn determine physiological fluxes, biomass partitioning and growth
231 rate at the organismal level [51,67]. In agreement, we have found that hybrids exhibit the same
232 allometric relationships as natural inbred accessions. This reinforces the idea of strong and
233 predictable constraints on phenotypic variation, specifically on variation of growth rate. Our
234 results further indicate that the shape of phenotypic non-linearity is important for the emergence
235 of heterosis [40]. For instance, a severe concavity is expected to be favourable to the emergence
236 of best-parent heterosis, as we have observed for fruit number. By contrast, phenotypic
237 convexity with a trait at lower integration level, for instance a trait related to organism
238 development and organization plan, could potentially lead to negative heterosis. This could
239 notably explain the high fraction of negative heterosis observed here for plant biomass, which is
240 consistent with previous findings in *A. thaliana* [13,54], although it contrasts with some others
241 [7,56]. It is also consistent with results in rice [17,68], an inbreeding crop species used for the
242 development of heterotic hybrids. In addition to phenotypic convexity, another explanation to
243 the emergence of negative heterosis is that we used crosses from strongly divergent accessions,
244 which might exhibit incompatibilities between defence-related genes [69-72]. Indeed,
245 autoimmunity has been repeatedly observed when crossing distant accessions of *A. thaliana* [64].
246 This is also consistent with recent findings, where positive heterosis for biomass seems to be
247 linked to a general repression of the immune system in *A. thaliana* hybrids [56].

248 The genetic bases of fitness-related traits such as growth rate and fruit production are
249 complex by nature, because these traits result from the effects of numerous genes acting on

250 different components of performance [26,73]. Inexpensive high-density genotype information
251 coupled with very detailed phenotyping provides a series of promising avenues for the genomic
252 prediction of heterosis [74]. In this context, the dominance hypothesis implicates, within certain
253 limits, a positive relationship between parental genetic distance and heterosis [28,29]. Results
254 from a range of species, however, do not conform with these expectations [7,13,34,75,76]. One
255 of the reasons could be the often small number of crosses and the relatively small range of
256 genetic distances analysed, with the latter holding true especially in cultivated species [7]. Our
257 results with 450 hybrids representing crosses between diverse *A. thaliana* populations pointed to
258 a positive but weak correlation between heterosis and parental genetic distance for growth rate,
259 but no such correlation for fruit number. This suggests that a genetic approach alone may not be
260 sufficient to accurately predict heterosis. Importantly, genetic distances calculated from the
261 genes enriched for the ones likely to have major effect on growth rate [27] did not improve our
262 ability to predict heterosis.

263 Wright's model of physiological dominance [38] was based on the non-linear
264 relationship that connects two traits at different levels of integration, enzyme concentration and
265 metabolic fluxes. In evolutionary biology, a trait is said to be "integrated" when it is closely
266 associated with the fitness of the organism, which in plants typically includes growth rate and
267 fruit number [77]. This leads to a pyramidal view of phenotypic integration [77,78], with fitness-
268 related traits at the top being under the control of several, less integrated traits (*e.g.*, biomass,
269 phenology, morphology), which themselves are under the control of many other traits (*e.g.*,
270 metabolic and physiological fluxes, cellular activity), and so on, until traits supposed to be
271 additive such as enzyme concentrations. Under the hypothesis of additivity of enzyme
272 concentration, Wright's model has been shown to accurately predict heterosis of metabolic
273 output in the chain of glycolysis [40]. In our study, we have shown that even when vegetative
274 dry mass exhibited important deviation from additivity, it could still be used to predict heterosis
275 of growth rate and fruit number based on parental allometric relationships. This finding
276 demonstrates the performance of the phenotypic approach relative to the purely genetic
277 approach to explain and predict heterosis. For instance, our method explained up to 75% of

278 heterosis amplitude for growth rate (Fig. 6B) while genetic distance at best explained only 7%
279 (Fig. 5A). Although our approach is not incompatible with the genetic models of molecular
280 dominance, it outperforms these in its ability to predict the direction and amplitude of heterosis
281 in the natural inbred species *A. thaliana*.

282 Evolutionary theory suggests that the intensity of inbreeding depression, and hence the
283 potential for heterosis, should increase with the level of phenotypic integration [79]. Consistent
284 with this, survival and fertility in animals are more sensitive to inbreeding depression than traits
285 at lower levels of integration such as size, biomass and gross morphology [80]. Our findings are
286 consistent with the idea that fitness-related traits exhibit more heterosis because they result from
287 the multiplicative effects of non-linear relationships at different levels of phenotypic integration
288 [48,73,81]. These non-linear relationships can have different curvatures and directions, which
289 might either amplify or cancel each other.

290 From a theoretical point of view, the predominance of outcrossing and the maintenance
291 of recessive alleles among organisms have been suggested to be directly linked to non-additive
292 inheritance and superior performance of hybrids [82]. Wright's initial model of physiological
293 dominance was proposed as a response to Fisher's idea that 'gene modifiers of dominance' must
294 exist and be selected to maximize the fitness of the heterozygote [83]. Wright [38], and later
295 Kacser and Burns [39], claimed that modifiers are not necessary, because non-linearity is an
296 intrinsic characteristic of metabolic networks. The same argument holds true for complex traits
297 such as growth rate and fruit number. The assumption of modifiers of dominance is based on
298 the unrealistic expectation of an intermediate phenotype in the heterozygote, while phenotypic
299 relationships are essentially non-linear [40,48,84]. A next step in getting to the physiological
300 root of heterosis will be the integration of other, more low-level traits, especially gene expression
301 and metabolite levels. We expect that this will improve the characterization of non-linear
302 relationships in multidimensional phenotypic space, and ultimately shed lights on the
303 physiological mechanisms at the origin of non-additive inheritance.

304 **Conclusions**

305 The development of a predictive approach for heterosis represents a long-term goal of modern
306 biology, especially in the applied framework of varietal selection in crops [85]. Our study
307 highlights the power of a geometric approach of trait-trait relationships for explaining heterosis
308 in two major components of plant productivity and yield. This in turn opens avenues for
309 targeting optimal crosses based on allometric relationships in parental lines. That trait variation
310 is similarly constrained in accessions and hybrids suggests that hybrids outperforming all
311 accessions in a specific trait are the exception, although they generally outperform their specific
312 parents. It is now time to test the phenotypic approach to heterosis in cultivated species, for
313 which the study of allometric relationships is a nascent research front [86–88].

314 **Material and Methods**

315 **Plant material**

316 451 accessions of *Arabidopsis thaliana* were phenotyped in 2014 at the Max Planck Institute for
317 Developmental Biology in Tübingen (MPI-Tübingen), Germany ($n = 2$, Exp1) [27,89]. 415
318 accessions were used as parental lines of the 450 hybrids phenotyped in a second experiment in
319 2015 at MPI-Tübingen ($n = 4$, Exp2). 342 accessions were used as female parent and 318 as
320 male parent. 407 accessions among the 451 total, and 369 accessions among the 415 parental,
321 had been genome sequenced as part of the 1001 Genomes project (<http://1001genomes.org/>)
322 [59]. Among the 415 parental accessions, 134 (32%) were used in a single cross, 166 (40%) were
323 used in two crosses, 63 (15%) in three crosses, and 52 (13%) in at least four crosses. To
324 overcome potential maternal effects, the same mother plants grown in the greenhouse in 2013
325 provided the seeds for both the inbreds (by self-fertilization) and the hybrids (by manual cross)
326 (Fig. S2A).

327 **Growth conditions**

328 We designed a hydroponic system where plants were cultivated on inorganic solid media
329 (rockwool) and all nutrients were provided through the watering solution. 4.6 cm (diameter) x 5
330 cm (depth) circular pots (Pöppelmann, Lohne, Germany) were filled with 3.6 cm x 3.6 cm x 4
331 cm depth rockwool cubes (Grodan cubes, Rockwool International, Denmark). Pots were
332 covered with a black foam disk with a 5-10 mm central circular opening. Seeds were sown in
333 individual pots, randomly distributed in trays of 30 pots each (Fig. S2B).

334 Before sowing, all seeds were surface-sterilized with 100% ethanol and frozen overnight
335 at -80 °C to kill any insect eggs. The rockwool cubes were placed in 75% strength nutrient
336 solution as described in ref. [90] in order to achieve full humidification and fertilization. After
337 sowing on the surface of the rockwool cubes, trays with 30 pots each were incubated for two
338 days in the dark at 4°C for stratification. Trays were transferred for 6 days to 23°C (8 h day
339 length) for germination. After 6 days, when most seedlings had two cotyledons, trays were
340 transferred to 4°C (8 h light) for 41 days of vernalization, in order to reduce the range of
341 flowering times among accessions. During germination and vernalization, all trays were
342 watered once a week with a 75% strength nutrient solution. After vernalization, when true
343 leaves had emerged on most individuals, plants were thinned to one individual per pot, and
344 trays were moved to the Raspberry Pi Automated Plant Analysis (RAPA) facility [89] (Fig.
345 S2B), set to 16°C, air humidity at 65%, and 12 h day length, with a PPFD of 125 to 175 $\mu\text{mol m}^{-2}$
346 s^{-1} provided by a 1:1 mixture of Cool White and Gro-Lux Wide Spectrum fluorescent lights
347 (Luxline plus F36W/840, Sylvania, Germany). Trays were randomly positioned in the room,
348 and watered every 1 to 3 days with 100% strength nutrient solution.

349 **Trait measurement**

350 Plants were grown and phenotyped using rigorously the same protocol, following
351 methodologies previously published for inbred accessions [27,89]. Plants were harvested at the
352 end of the life cycle when the first fruits were senescing. Rosettes were separated from roots and
353 reproductive parts, dried at 65° C for three days and weighed. Plant lifespan (d) was measured

354 as the duration between the appearance of the two first leaves after vernalization and the end of
355 the life cycle [89]. Growth rate (mg d^{-1}) was calculated as the ratio of final rosette dry mass over
356 plant lifespan. Inflorescences were photographed with a high-resolution, 16.6 megapixel SLR
357 camera (Canon EOS-1, Canon Inc., Japan), and analysed with ImageJ [91] to estimate the
358 number of fruits through 2D image skeletonization, following published protocols [27,89].

359 The RAPA system was used for daily imaging using 192 micro-cameras (OmniVision
360 OV5647), which simultaneously acquired 6 daily top-view 5-megapixel images for each tray
361 during the first 25 days after vernalization. We used a published method to estimate plant dry
362 mass during ontogeny from top-view rosette pictures [27,89]. From fitted sigmoid growth curves
363 on all individuals, we calculated inflection point (d) at which daily growth was maximal and
364 started to decrease. We used rosette dry mass (mg) at the inflection point as measurement of
365 vegetative dry mass M (mg). In total, trait values for plant lifespan, growth rate and vegetative
366 dry mass were available on 451 inbreds and 447 hybrids, and fruit number on 441 inbreds and
367 449 hybrids (Table S1).

368 To correct for potential biases between the two experiments performed, 16 accessions
369 phenotyped in Exp1 were also included in Exp2. Among all traits measured, only plant lifespan
370 exhibited a significant difference between the two experiments ($P = 0.03$). We thus corrected
371 trait values in Exp2 with the following equation: $\text{Lifespan}_{\text{corrected}} = -37 + 1.8 * \text{Lifespan}_{\text{observed}}$.

372 **Measurement of heterosis**

373 First, hybrid phenotypic classes were categorized by comparing trait distribution in hybrids to
374 mean and best (or worst) parental values. We used 95% confidence intervals (CI) among the
375 four hybrid replicates of each F1 for categorization, as follows: (i) below worst-parent if CI of
376 the trait measured in the hybrid was strictly inferior to the minimum trait value among the two
377 parental genotypes; (ii) below mean-parent if the hybrid CI was strictly inferior to mean parental
378 value but overlapped with minimum parent; (iii) above mean-parent if the hybrid CI was strictly

379 superior to mean parental value (but not best parent); (iv) above best-parent if the hybrid CI was
380 strictly superior to maximum parental value.

381 Secondly, two metrics commonly used in the literature to quantify heterosis [13] were
382 calculated in the present study for all traits, Y , *i.e.* growth rate and fruit number:

383 1) Mid-Parent Heterosis (MPH), the deviation of observed hybrid value scaled to observed
384 mean parental value:

$$385 \quad \text{MPH} = (Y_{1 \times 2} - \text{mean}(Y_1, Y_2) / \text{mean}(Y_1, Y_2)$$

386 2) Best-Parent Heterosis (BPH), the deviation of observed hybrid value scaled to observed
387 best parental value:

$$388 \quad \text{BPH} = (Y_{1 \times 2} - \max(Y_1, Y_2) / \max(Y_1, Y_2)$$

389 **Allometric relationships and measurement of phenotypic non-linearity**

390 The allometric equations were fitted by the non-linear least-squares method using the *nls*
391 function in R [92]. Based on allometric theory [27,43,44,52], we chose a power-law equation for
392 growth rate with a mass-corrected allometric exponent ($g(M)$ in Table 1). The corrected
393 exponent corresponds to the derivative of the quadratic function obtained after logarithmic
394 transformation of the allometric relationship [27,43,52]. In our data, biomass-correction of the
395 exponent improved the fitting of the allometric relationship of growth rate ($\Delta\text{AIC} = 83$). For the
396 allometric relationship of fruits number, we chose an inverse polynomial equation ($f(M)$ in Table
397 1). 95% CIs of the fitted coefficients were estimated with the *confint* function in R.

398 We used the allometric equations fitted on the accessions to predict the phenotype of the
399 hybrids from vegetative dry mass M . We first estimated growth rate of both parents and hybrids
400 ($g(M_1)$, $g(M_2)$, $g(M_{1 \times 2})$ with M_1 and M_2 corresponding to vegetative dry mass of worst and best
401 parent, respectively, and $M_{1 \times 2}$ to hybrid dry mass). We then predicted mean parental value of
402 growth rate as $\text{MP}_{\text{pred}} = (g(M_1) + g(M_2))/2$, and best parental value as $\text{BP}_{\text{pred}} = \max(g(M_1), g(M_2))$.
403 Finally, we predicted the phenotypic non-linearity as:

404 1) the deviation of predicted hybrid value scaled to predicted mean parental value (Fig. 6A):

$$405 \quad \text{PNL}_{\text{MP}} = (g(M_{1 \times 2}) - \text{MP}_{\text{pred}}) / \text{MP}_{\text{pred}}$$

406 2) the deviation of predicted hybrid value scaled to predicted best parental value (Fig. 6C):

$$407 \quad \text{PNL}_{\text{BP}} = (g(M_{1 \times 2}) - \text{BP}_{\text{pred}}) / \text{BP}_{\text{pred}}$$

408 Both PNL_{MP} and PNL_{BP} were calculated for growth rate, and we performed similarly for
409 estimating the non-linearity of fruit number (Fig. 6C).

410 **Genetic, geographic and phenotypic distances**

411 407 accessions out of the 451 phenotyped here have been genome sequenced
412 (<http://1001genomes.org/>) [59]. Using vcftools [93], the 12,883,854 single nucleotide
413 polymorphisms (SNPs) were first filtered to retain those where minor allele frequency was above
414 5%, with a genotyping rate above 85% across all accessions. This resulted in 391,016 SNPs. We
415 used PLINK v1.9 [94] to estimate pairwise genetic distances as the number of alleles that
416 differed between pairs of accessions (*--distance* function), after \log_{10} -transformation. We also
417 measured genetic distances using the 1% SNPs with the strongest (positive and negative) effects
418 on growth rate or fruit number from a previously published study [27]. SNPs effects on each
419 trait were estimated using a polygenic GWA model implemented in GEMMA ('Bayesian
420 Sparse Linear Mixed Model', BSLMM) [95].

421 Pairwise geographic distances between accessions were estimated from their longitude-
422 latitude coordinates [59] and the *distm* function of the *geosphere* package in R. For the
423 calculation of pairwise phenotypic distances, we first used Euclidean distance among all traits
424 measured in inbreds (vegetative dry mass, lifespan, growth rate, fruit number; used for Fig. 1).
425 We also used absolute difference in vegetative dry mass between parents for comparing the
426 contribution of genetic and phenotypic distances to heterosis in Fig. 5.

427 **Statistical analyses**

428 Pearson's method was used for correlation analyses (*cor.test* function in R). Linear regressions
429 were drawn with Standard Major Axis (SMA, 'smatr' package in R). The effect of the
430 experiment on trait values was tested on the 16 accessions common to Exp1 and Exp2, using
431 two-way ANOVA with genotype and experiment as interacting factors. 95% Confidence
432 Interval (CI) was calculated as 1.96*Standard Deviation. To calculate the proportion of
433 phenotypic variance associated with genotypic (G) effects (a measure of broad-sense heritability,
434 $H^2 = \text{var}(G) / (\text{var}(G) + \text{residuals})$), we fitted a mixed model (*lmer* function in R) as $Y =$
435 Genotype + residuals, where Y is trait value and Genotype is used as random factor. Statistical
436 analyses were conducted in R v3.2.3 [92].

437 **Availability of data and material**

438 The datasets supporting the conclusions of this article are included within the article and its
439 additional files. Codes are available on Github
440 (<https://github.com/fvasseur/AnalysisHeterosis>), and phenotypic data are available on the
441 Dryad repository (<https://datadryad.org/review?doi=doi:10.5061/dryad.4978rp4>) [96].
442 Correspondence and requests for materials should be addressed to weigel@weigelworld.org or
443 franc.vasseur@gmail.com.

444 **Competing interests**

445 The authors declare no competing financial interests.

446 **Author Contributions**

447 FV, CV and DW formulated hypotheses and designed the study. FV performed the experiments
448 and extracted the data. FV, LF and DV performed statistical analyses. All authors interpreted
449 the results and wrote the paper.

450 **Acknowledgements**

451 We thank Christine Dillmann, Alain Charcosset, Olivier Loudet and Thomas Lenormand for
452 helpful comments about statistical modelling and data interpretation.

453 **References**

- 454 1. Melchinger AE. Genetic diversity and heterosis. In: Coors JG, Pandey S, editors. The
455 genetics and exploitation of heterosis in crops. American Society of Agronomy, Crop
456 Science Society of America; 1999. pp. 99–118.
- 457 2. Fu D, Xiao M, Hayward A, Jiang G, Zhu L, Zhou Q, et al. What is crop heterosis: new
458 insights into an old topic. *J Appl Genet*. 2015;56: 1–13.
- 459 3. Charlesworth D, Willis JH. The genetics of inbreeding depression. *Nat Rev Genet*.
460 2009;10: 783–796.
- 461 4. Zhou G, Chen Y, Yao W, Zhang C, Xie W, Hua J, et al. Genetic composition of yield
462 heterosis in an elite rice hybrid. *Proc Natl Acad Sci U S A*. 2012;109: 15847–15852.
- 463 5. Garcia AAF, Wang S, Melchinger AE, Zeng Z-B. Quantitative trait loci mapping and the
464 genetic basis of heterosis in maize and rice. *Genetics*. 2008;180: 1707–1724.
- 465 6. Flint-Garcia SA, Buckler ES, Tiffin P, Ersoz E, Springer NM. Heterosis is prevalent for
466 multiple traits in diverse maize germplasm. *PLoS One*. 2009;4: e7433.
- 467 7. Barth S, Busimi AK, Friedrich Utz H, Melchinger AE. Heterosis for biomass yield and
468 related traits in five hybrids of *Arabidopsis thaliana* L. *Heynh*. *Heredity*. 2003;91: 36–42.
- 469 8. Barbosa-Neto JF, Sorrells ME, Cisar G. Prediction of heterosis in wheat using coefficient
470 of parentage and RFLP-based estimates of genetic relationship. *Genome*. 1996;39: 1142–
471 1149.
- 472 9. Krieger U, Lippman ZB, Zamir D. The flowering gene SINGLE FLOWER TRUSS drives
473 heterosis for yield in tomato. *Nat Genet*. 2010;42: 459–463.
- 474 10. Birchler JA, Yao H, Chudalayandi S. Unraveling the genetic basis of hybrid vigor. *Proc*
475 *Natl Acad Sci U S A*. 2006;103: 12957–12958.
- 476 11. Birchler JA, Auger DL, Riddle NC. In search of the molecular basis of heterosis. *Plant*
477 *Cell*. 2003;15: 2236–2239.
- 478 12. Chen ZJ. Genomic and epigenetic insights into the molecular bases of heterosis. *Nat Rev*
479 *Genet*. 2013;14: 471–482.

- 480 13. Seymour DK, Chae E, Grimm DG, Martín Pizarro C, Habring-Müller A, Vasseur F, et al.
481 Genetic architecture of nonadditive inheritance in *Arabidopsis thaliana* hybrids. *Proc Natl*
482 *Acad Sci U S A*. 2016;113: E7317–E7326.
- 483 14. Hua J, Xing Y, Wu W, Xu C, Sun X, Yu S, et al. Single-locus heterotic effects and
484 dominance by dominance interactions can adequately explain the genetic basis of heterosis
485 in an elite rice hybrid. *Proc Natl Acad Sci U S A*. 2003;100: 2574–2579.
- 486 15. Melchinger AE, Piepho H-P, Utz HF, Muminovic J, Wegenast T, Törjék O, et al. Genetic
487 basis of heterosis for growth-related traits in *Arabidopsis* investigated by testcross progenies
488 of near-isogenic lines reveals a significant role of epistasis. *Genetics*. 2007;177: 1827–1837.
- 489 16. Li L, Lu K, Chen Z, Mu T, Hu Z, Li X. Dominance, overdominance and epistasis
490 condition the heterosis in two heterotic rice hybrids. *Genetics*. 2008;180: 1725–1742.
- 491 17. Li ZK, Luo LJ, Mei HW, Wang DL, Shu QY, Tabien R, et al. Overdominant epistatic loci
492 are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and
493 grain yield. *Genetics*. 2001;158: 1737–1753.
- 494 18. Xiao J, Li J, Yuan L, Tanksley SD. Dominance is the major genetic basis of heterosis in
495 rice as revealed by QTL analysis using molecular markers. *Genetics*. 1995;140: 745–754.
- 496 19. Kaeppeler S. Heterosis: many genes, many mechanisms—end the search for an
497 undiscovered unifying theory. *ISRN Botany*. 2012;2012.
- 498 20. Lippman ZB, Zamir D. Heterosis: revisiting the magic. *Trends Genet*. 2007;23: 60–66.
- 499 21. Kusterer B, Muminovic J, Utz HF, Piepho H-P, Barth S, Heckenberger M, et al. Analysis
500 of a triple testcross design with recombinant inbred lines reveals a significant role of
501 epistasis in heterosis for biomass-related traits in *Arabidopsis*. *Genetics*. 2007;175: 2009–
502 2017.
- 503 22. Groszmann M, Greaves IK, Fujimoto R, Peacock WJ, Dennis ES. The role of epigenetics
504 in hybrid vigour. *Trends Genet*. 2013;29: 684–690.
- 505 23. Larièpe A, Mangin B, Jasson S, Combes V, Dumas F, Jamin P, et al. The genetic basis of
506 heterosis: multiparental quantitative trait loci mapping reveals contrasted levels of apparent
507 overdominance among traits of agronomical interest in maize (*Zea mays L.*). *Genetics*.
508 2012;190: 795–811.
- 509 24. Rockman MV. The QTN program and the alleles that matter for evolution: all that's gold
510 does not glitter. *Evolution*. 2012;66: 1–17.
- 511 25. Kroymann J, Mitchell-Olds T. Epistasis and balanced polymorphism influencing complex
512 trait variation. *Nature*. 2005;435: 95–98.
- 513 26. Holland JB. Genetic architecture of complex traits in plants. *Curr Opin Plant Biol*. 2007;10:
514 156–161.
- 515 27. Vasseur F, Exposito-Alonso M, Ayala-Garay OJ, Wang G, Enquist BJ, Vile D, et al.

- 516 Adaptive diversification of growth allometry in the plant *Arabidopsis thaliana*. Proc Natl
517 Acad Sci U S A. 2018; 201709141.
- 518 28. Burstin J, Charcosset A. Relationship between phenotypic and marker distances:
519 theoretical and experimental investigations. Heredity. 1997;79.
- 520 29. Charcosset A, Lefort-Buson M, Gallais A. Relationship between heterosis and
521 heterozygosity at marker loci: a theoretical computation. Theor Appl Genet. 1991;81: 571–
522 575.
- 523 30. Charcosset A, Essioux L. The effect of population structure on the relationship between
524 heterosis and heterozygosity at marker loci. Theor Appl Genet. 1994;89: 336–343.
- 525 31. Moll RH, Lonquist JH, Fortuno JV, Johnson EC. The relationship of heterosis and
526 genetic divergence in maize. Genetics. 1965;52: 139–144.
- 527 32. Xiao J, Li J, Yuan L, McCouch SR, Tanksley SD. Genetic diversity and its relationship to
528 hybrid performance and heterosis in rice as revealed by PCR-based markers. Theor Appl
529 Genet. 1996;92: 637–643.
- 530 33. Betran FJ, Ribaut JM, Beck D, De Leon DG. Genetic diversity, specific combining ability,
531 and heterosis in tropical maize under stress and nonstress environments. Crop Sci. 2003;43:
532 797–806.
- 533 34. Palacio-Lopez K, Keller SR, Molofsky J. Genomic admixture between locally adapted
534 populations of *Arabidopsis thaliana* (Mouse ear cress): Evidence of optimal genetic
535 outcrossing distance. J Hered. 2017;109(1): 38–46.
- 536 35. Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, et al. Whole-genome
537 sequencing of multiple *Arabidopsis thaliana* populations. Nat Genet. 2011;43: 956–963.
- 538 36. Birchler JA, Yao H, Chudalayandi S, Vaiman D, Veitia RA. Heterosis. Plant Cell. 2010;22:
539 2105–2112.
- 540 37. Goff SA. A unifying theory for general multigenic heterosis: energy efficiency, protein
541 metabolism, and implications for molecular breeding. New Phytol. 2011;189: 923–937.
- 542 38. Wright S. Physiological and evolutionary theories of dominance. Am Nat. 1934;68: 24–53.
- 543 39. Kacser H, Burns JA. The molecular basis of dominance. Genetics. 1981;97: 639–666.
- 544 40. Fiévet J, Nidelet T, Dillmann C, de Vienne D. Heterosis is a systemic property emerging
545 from nonlinear genotype-phenotype relationships: evidence from in vitro genetics and
546 computer simulations. Front Genet. 2018;9: 159.
- 547 41. Fiévet JB, Dillmann C, de Vienne D. Systemic properties of metabolic networks lead to an
548 epistasis-based model for heterosis. Theor Appl Genet. 2010;120: 463–473.
- 549 42. Niklas KJ. Plant allometry: the scaling of form and process. University of Chicago Press;
550 1994.

- 551 43. Kolokotronis T, Van Savage, Deeds EJ, Fontana W. Curvature in metabolic scaling.
552 Nature. 2010;464: 753–756.
- 553 44. Price CA, Enquist BJ, Savage VM. A general model for allometric covariation in botanical
554 form and function. Proc Natl Acad Sci U S A. 2007;104: 13204–13209.
- 555 45. West GB, Brown JH, Enquist BJ. The fourth dimension of life: fractal geometry and
556 allometric scaling of organisms. Science. 1999;284: 1677–1679.
- 557 46. Brown JH, Gillooly JF, Allen AP, Savage VM, West GB. Toward a metabolic theory of
558 ecology. Ecology. 2004;85: 1771–1789.
- 559 47. Enquist BJ, Niklas KJ. Global allocation rules for patterns of biomass partitioning in seed
560 plants. Science. 2002;295: 1517–1520.
- 561 48. Kerkhoff AJ, Enquist BJ. Multiplicative by nature: why logarithmic transformation is
562 necessary in allometry. J Theor Biol. 2009;257: 519–521.
- 563 49. Savage VM, Gillooly JF, Woodruff WH, West GB, Allen AP, Enquist BJ, et al. The
564 predominance of quarter-power scaling in biology. Funct Ecol. 2004;18: 257–282.
- 565 50. Savage VM, Bentley LP, Enquist BJ, Sperry JS, Smith DD, Reich PB, et al. Hydraulic
566 trade-offs and space filling enable better predictions of vascular structure and function in
567 plants. Proc Natl Acad Sci U S A. 2010;107: 22722–22727.
- 568 51. Enquist BJ, Kerkhoff AJ, Stark SC, Swenson NG, McCarthy MC, Price CA. A general
569 integrative model for scaling plant growth, carbon flux, and functional trait spectra. Nature.
570 2007;449: 218–222.
- 571 52. Vasseur F, Violle C, Enquist BJ, Granier C, Vile D. A common genetic basis to the origin
572 of the leaf economics spectrum and metabolic scaling allometry. Ecol Lett. 2012;15: 1149–
573 1157.
- 574 53. Somerville C, Somerville S. Plant functional genomics. Science. 1999;285: 380–383.
- 575 54. Meyer RC, Törjék O, Becher M, Altmann T. Heterosis of biomass production in
576 Arabidopsis. Establishment during early development. Plant Physiol. 2004;134: 1813–1823.
- 577 55. Meyer RC, Kusterer B, Lisek J, Steinfath M, Becher M, Scharf H, et al. QTL analysis of
578 early stage heterosis for biomass in Arabidopsis. Theor Appl Genet. 2010;120: 227–237.
- 579 56. Yang M, Wang X, Ren D, Huang H, Xu M, He G, et al. Genomic architecture of biomass
580 heterosis in Arabidopsis. Proc Natl Acad Sci U S A. 2017;114(30): 8101–8106.
- 581 57. Abbott RJ, Gomes MF. Population genetic structure and outcrossing rate of *Arabidopsis*
582 *thaliana* (L.) Heynh. Heredity. 1989;62: 411.
- 583 58. Bomblies K, Yant L, Laitinen RA, Kim S-T, Hollister JD, Warthmann N, et al. Local-scale
584 patterns of genetic variability, outcrossing, and spatial structure in natural stands of
585 *Arabidopsis thaliana*. PLoS Genet. 2010;6: e1000890.

- 586 59. 1001 Genomes Consortium. 1135 sequenced natural inbred lines reveal the global pattern
587 of polymorphism in *Arabidopsis thaliana*. *Cell*. 2016;166: 481–491.
- 588 60. Koornneef M, Reymond M, Alonso-Blanco C. Natural variation in *Arabidopsis thaliana*. In:
589 Schmidt R, Bancroft I, editors. *Genetics and genomics of the Brassicaceae*. Springer, New
590 York, NY; 2011. pp. 123–151.
- 591 61. Mitchell-Olds T, Schmitt J. Genetic mechanisms and evolutionary significance of natural
592 variation in *Arabidopsis*. *Nature*. 2006;441: 947–952.
- 593 62. Bergelson J, Roux F. Towards identifying genes underlying ecologically relevant traits in
594 *Arabidopsis thaliana*. *Nat Rev Genet*. 2010;11: 867–879.
- 595 63. Weigel D. Natural variation in *Arabidopsis*: from molecular genetics to ecological
596 genomics. *Plant Physiol*. 2012;158: 2–22.
- 597 64. Chae E, Bomblies K, Kim S-T, Karelina D, Zaidem M, Ossowski S, et al. Species-wide
598 genetic incompatibility analysis identifies immune genes as hot spots of deleterious
599 epistasis. *Cell*. 2014;159: 1341–1351.
- 600 65. Bonser SP, Aarssen LW. Allometry and plasticity of meristem allocation throughout
601 development in *Arabidopsis thaliana*. *J Ecol*. 2001;89: 72–79.
- 602 66. West GB, Brown JH, Enquist BJ. A general model for the structure and allometry of plant
603 vascular systems. *Nature*. 1999;400: 664–667.
- 604 67. Niklas KJ, Enquist BJ. Invariant scaling relationships for interspecific plant biomass
605 production rates and body size. *Proc Natl Acad Sci U S A*. 2001;98: 2922–2927.
- 606 68. Xiang C, Zhang H, Wang H, Wei S, Fu B, Xia J, et al. Dissection of heterosis for yield and
607 related traits using populations derived from introgression lines in rice. *The Crop Journal*.
608 2016;4: 468–478.
- 609 69. Bomblies K, Lempe J, Epple P, Warthmann N, Lanz C, Dangl JL, et al. Autoimmune
610 response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in
611 plants. *PLoS Biol*. 2007;5: 1962–1972.
- 612 70. Świadek M, Proost S, Sieh D, Yu J, Todesco M, Jorzig C, et al. Novel allelic variants in
613 *ACD6* cause hybrid necrosis in local collection of *Arabidopsis thaliana*. *New Phytol*.
614 2017;213: 900–915.
- 615 71. Todesco M, Kim S-T, Chae E, Bomblies K, Zaidem M, Smith LM, et al. Activation of the
616 *Arabidopsis thaliana* immune system by combinations of common *ACD6* alleles. *PLoS*
617 *Genet*. 2014;10: e1004459.
- 618 72. Todesco M, Balasubramanian S, Hu TT, Traw MB, Horton M, Epple P, et al. Natural
619 allelic variation underlying a major fitness trade-off in *Arabidopsis thaliana*. *Nature*.
620 2010;465: 632–636.
- 621 73. Grafius JE. The complex trait as a geometric construct. *Heredity*. 1961;16: 225.

- 622 74. Heffner EL, Sorrells ME, Jannink JL. Genomic selection for crop improvement. *Crop Sci.*
623 2009;49(1): 1–12.
- 624 75. Liu X, Ishiki K, Wang W. Identification of AFLP markers favorable to heterosis in hybrid
625 rice. *Breed Sci.* 2002;52: 201–206.
- 626 76. Barbosa AMM, Geraldi IO, Benchimol LL, Garcia AAF, Souza CL, Souza AP.
627 Relationship of intra- and interpopulation tropical maize single cross hybrid performance
628 and genetic distances computed from AFLP and SSR markers. *Euphytica.* 2003;130: 87–
629 99.
- 630 77. Arnold SJ. Morphology, performance and fitness. *Integr Comp Biol.* 1983;23: 347–361.
- 631 78. Violle C, Navas ML, Vile D, Kazakou E, Fortunel C, Hummel I, et al. Let the concept of
632 trait be functional! *Oikos.* 2007;116: 882–892.
- 633 79. Falconer DS. *Introduction to Quantitative Genetics* Ed. 3. Harlow, Essex, UK/New York:
634 Longmans Green/John Wiley & Sons; 1989.
- 635 80. DeRose MA, Roff DA. A comparison of inbreeding depression in life-history and
636 morphological traits in animals. *Evolution.* 1999;53: 1288–1292.
- 637 81. Rendel JM. The relationship between gene and phenotype. *J Theor Biol.* 1962;2: 296–308.
- 638 82. Charlesworth D, Charlesworth B. Inbreeding depression and its evolutionary
639 consequences. *Annu Rev Ecol Syst.* 1987;18: 237–268.
- 640 83. Fisher RA. *The genetical theory of natural selection.* Oxford, England: Clarendon Press;
641 1930.
- 642 84. Paine CE, Marthews TR, Vogt DR, Purves D, Rees M, Hector A, et al. How to fit
643 nonlinear plant growth models and calculate growth rates: an update for ecologists.
644 *Methods Ecol Evol.* 2012;3: 245–256.
- 645 85. Xu S, Zhu D, Zhang Q. Predicting hybrid performance in rice using genomic best linear
646 unbiased prediction. *Proc Natl Acad Sci U S A.* 2014;111: 12456–12461.
- 647 86. Milla R, Osborne CP, Turcotte MM, Violle C. Plant domestication through an ecological
648 lens. *Trends Ecol Evol.* 2015;30: 463–469.
- 649 87. Deng J, Zuo W, Wang Z, Fan Z, Ji M, Wang G, et al. Insights into plant size-density
650 relationships from models and agricultural crops. *Proc Natl Acad Sci U S A.* 2012;109:
651 8600–8605.
- 652 88. Deng J, Ran J, Wang Z, Fan Z, Wang G, Ji M, et al. Models and tests of optimal density
653 and maximal yield for crop plants. *Proc Natl Acad Sci U S A.* 2012;109: 15823–15828.
- 654 89. Vasseur F, Wang G, Bresson J, Schwab R, Weigel D. Image-based methods for
655 phenotyping growth dynamics and fitness in *Arabidopsis thaliana*. *bioRxiv.* 2018. p. 208512.
- 656 90. Conn SJ, Hocking B, Dayod M, Xu B, Athman A, Henderson S, et al. Protocol: optimising

- 657 hydroponic growth systems for nutritional and physiological analysis of *Arabidopsis thaliana*
658 and other plants. *Plant Methods*. 2013;9: 4.
- 659 91. Rasband WS. 1997–2011. ImageJ. US National Institutes of Health, Bethesda, MD, USA:
660 <http://imagej.nih.gov/ij>. 2011.
- 661 92. Team RC. R: A language and environment for statistical computing. R foundation for
662 statistical computing, Vienna, Austria. 2014. ISBN 3-900051-07-0.
- 663 93. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant
664 call format and VCFtools. *Bioinformatics*. 2011;27: 2156–2158.
- 665 94. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a
666 tool set for whole-genome association and population-based linkage analyses. *Am J Hum*
667 *Genet*. 2007;81: 559–575.
- 668 95. Zhou X, Stephens M. Efficient multivariate linear mixed model algorithms for genome-
669 wide association studies. *Nat Methods*. 2014;11: 407–409.
- 670 96. Greenberg J, White HC, Carrier S, Scherle R. A metadata best practice for a scientific data
671 repository. *J Libr Metadata*. 2009;9: 194–212.
- 672

673 **Table**

674 **Table 1. Coefficients of fitted allometric relationships.**

Equation		Fitted coefficients [95% CI]		
		<i>a</i>	<i>b</i>	<i>c</i>
$g(M) = aM^{b+Log(M)}$	Inbreds	0.022 [0.011;0.043]	1.479 [1.248;1.717]	-0.162 [-0.208;-0.118]
	Hybrids	0.019 [0.010;0.037]	1.471 [1.255;1.694]	-0.154 [-0.196;-0.113]
$f(M) = \frac{M}{a + bM + cM^2}$	Inbreds	0.154 [0.107;0.210]	0.0022 [0.0015;0.0029]	1.291 10⁻⁵ [1.085 10 ⁻⁵ ;1.523 10 ⁻⁵]
	Hybrids	0.119 [0.096;0.144]	0.0025 [0.0021;0.0030]	1.456 10⁻⁵ [1.301 10 ⁻⁵ ;1.622 10 ⁻⁵]

675

676 **Figure legends**

677 **Figure 1. Wright's model of physiological dominance.** The metabolic flux along the y axis
678 exhibits heterosis as a result of non-linear relationship with the concentration of enzymes along
679 the x axis. This is expected for hybrids (genotype Aa) even if the concentration of enzymes is
680 purely additive, as shown here (\bar{X} is mean enzyme concentration value between parent 1 (aa)
681 and parent 2 (AA)). Red line represents Michaelis-Menten kinetics of enzyme activity ($k(X)$).

682 **Figure 2. Diversity of genetic crosses used.** (A) Origin of 451 natural inbred accessions (red
683 dots), and the 450 hybrids between them (blue connecting lines). (B) Relationship between
684 pairwise genetic distances and pairwise geographic distances. (C) Relationship between pairwise
685 genetic distances and pairwise phenotypic distances (Euclidean distances). Distance-distance
686 relationships are represented between all possible pairs of the 451 accessions (101,475
687 combinations, grey dots) and the crosses used for the 451 phenotyped hybrids (blue dots).

688 **Figure 3. Distributions of traits and proportion of heterosis.** (A) Vegetative dry mass M in
689 inbreds (red) and hybrids (blue). Pie chart indicates proportion of crosses without and with
690 significant heterosis. (B) Plant lifespan. (C) Growth rate. (D) Fruit number.

691 **Figure 4. Allometric relationships of growth rate and fruit number.** (A) Allometric
692 relationship between growth rate and vegetative plant dry mass, M , fitted with a power-law
693 function with mass-corrected exponent ($g(M)$). Equation fitted separately on 451 inbreds (red
694 dots, solid line) and 447 hybrids (blue dots, dashed line). (B) Allometric relationship between
695 fruit number and M , fitted with an inverse polynomial function ($f(M)$). Equation fitted separately
696 on 441 inbreds (red dots, solid line) and 449 hybrids (blue dots, dashed line).

697 **Figure 5. Correlation of heterosis with genetic and phenotypic distances.** (A) Correlation
698 between pairwise genetic distances between parental accessions and observed heterosis of
699 growth rate in hybrids ($n = 368$). (B) Correlation between pairwise phenotypic distances
700 between parental accessions (absolute difference in M) and observed heterosis of growth rate in
701 hybrids ($n = 447$). (C) Pairwise genetic distances versus heterosis of fruit number ($n = 368$). (D)
702 Pairwise phenotypic distances versus heterosis of fruit number ($n = 449$). MPH: Mid-Parent
703 Heterosis (light blue), BPH: Best-parent heterosis (dark blue). r^2 are Pearson's coefficients of

704 correlation (NS: Non-significant, ***: $P < 0.001$). Regressions lines are SMA, drawn when
705 significant (solid lines). Red dashed lines represent zero axis.

706 **Figure 6. Prediction of heterosis with phenotypic non-linearity.** (A) Allometric relationship of
707 growth rate ($g(M)$, black solid line) fitted in inbreds, with an example of two parental accessions
708 (Udul 4-9 and Dra2-1, in red), and their hybrid (in blue). Phenotypic non-linearity (PNL, blue
709 arrows) was calculated relative to mid-parent value (MP_{pred}): $PNL_{MP} = (g(M_{1x2}) - MP_{pred}) /$
710 MP_{pred} , and relative to best-parent value (BP_{pred}): $PNL_{BP} = (g(M_{1x2}) - BP_{pred}) / BP_{pred}$. (B)
711 Correlation between PNL and heterosis in all hybrids ($n = 447$), both relative to mid- and best-
712 parent value (MPH in light blue and BPH in dark blue, respectively). (C) Allometric relationship
713 of fruit number ($f(M)$, black solid line) fitted in inbreds, with an example of two parental
714 accessions (Cdm-0 and Monte-1, in red), and their hybrid (in blue). PNL was calculated relative
715 to mid-parent value (PNL_{MP}) and relative to best-parent value (PNL_{BP}). (D) Correlation between
716 PNL and heterosis in all hybrids ($n = 447$), both relative to MPH (light blue) and BPH (dark
717 blue). Red dashed lines represent zero axes. Black dashed line represents 1:1 line. Grey dots are
718 all hybrid individuals. Pictures of plants were taken at harvesting, at the end of reproduction. r^2
719 are Pearson's coefficients of correlation (NS: Non-significant, ***: $P < 0.001$).

720 **Additional Files**

721 **Additional Table S1. Traits measured in inbred accessions and hybrids of *A. thaliana***

722 **Additional Table S2. Summary statistics of traits measured.**

723 **Additional Table S3. Correlation of heterosis with genetic and phenotypic distances.**

724 **Additional Figure S1. Correlation between predicted and observed hybrid phenotype. (A)**

725 Hybrid growth rate (mg d^{-1}) predicted by phenotypic non-linearity (brown dots) and
726 genetic additivity (mean growth rate between parents, yellow dots), and compared to
727 observed hybrid value ($n = 447$). **(B)** Hybrid fruit number predicted by phenotypic non-
728 linearity (brown dots) and genetic additivity (mean fruit number between parents, yellow
729 dots), and compared to observed hybrid value ($n = 449$). r^2 are Pearson's coefficients of
730 correlation (NS: Non-significant, ***: $P < 0.001$). Dashed line represents 1:1 line.

731 **Additional Figure S2. Crossing and phenotyping conditions. (A)** Seed production experiment

732 performed in 2013 at MPI-Tübingen (Germany). After manual crossing or self-
733 fertilization, mother flowers were isolated in a small paper bags until fruit ripening.
734 Seeds for accessions and hybrids used in this study came from the same mother plants.
735 **(B)** RAPA growth chamber at MPI-Tübingen with trays of accessions phenotyped
736 during Exp1 in 2014.

Figure 1.

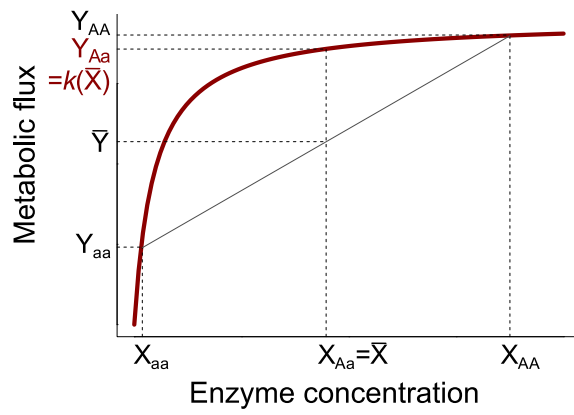


Figure 2.

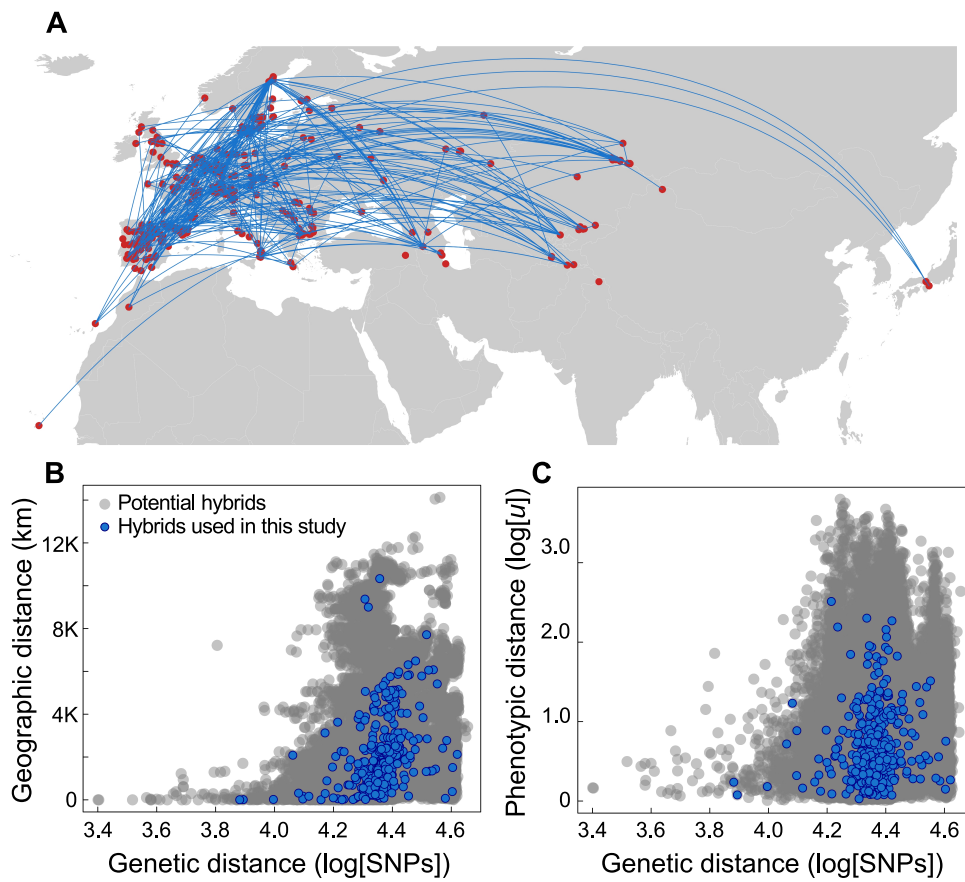


Figure 3.

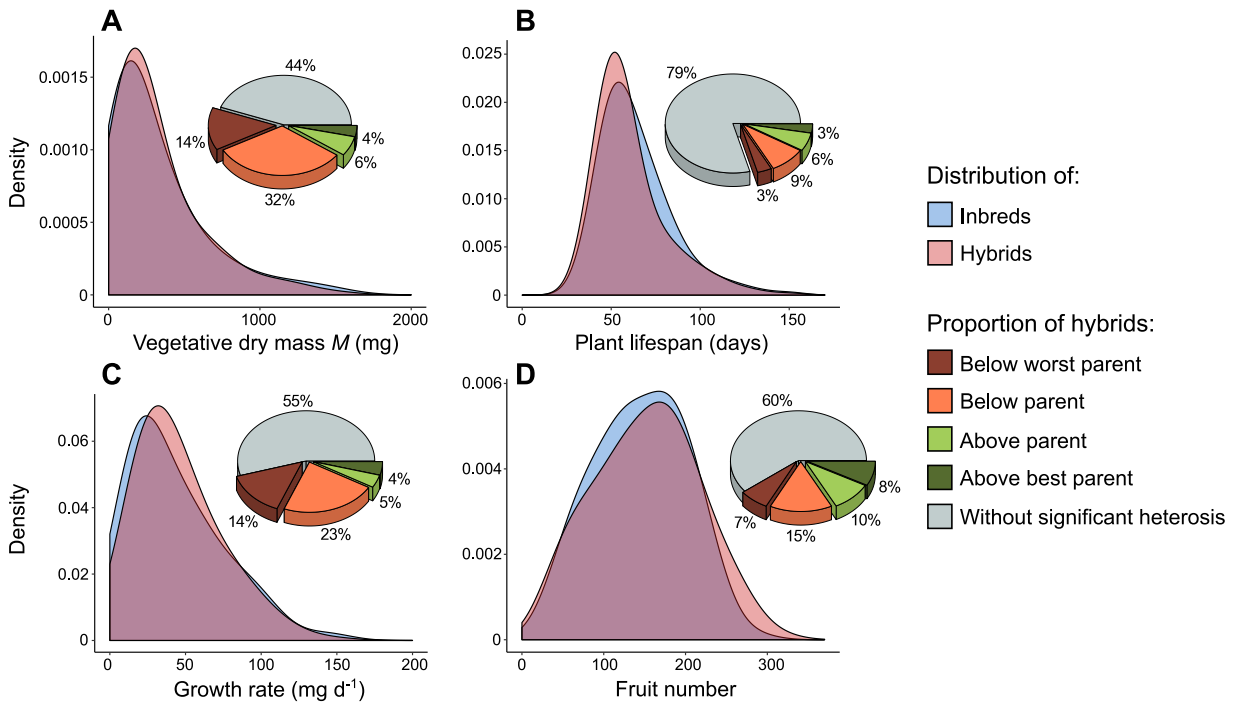


Figure 4.

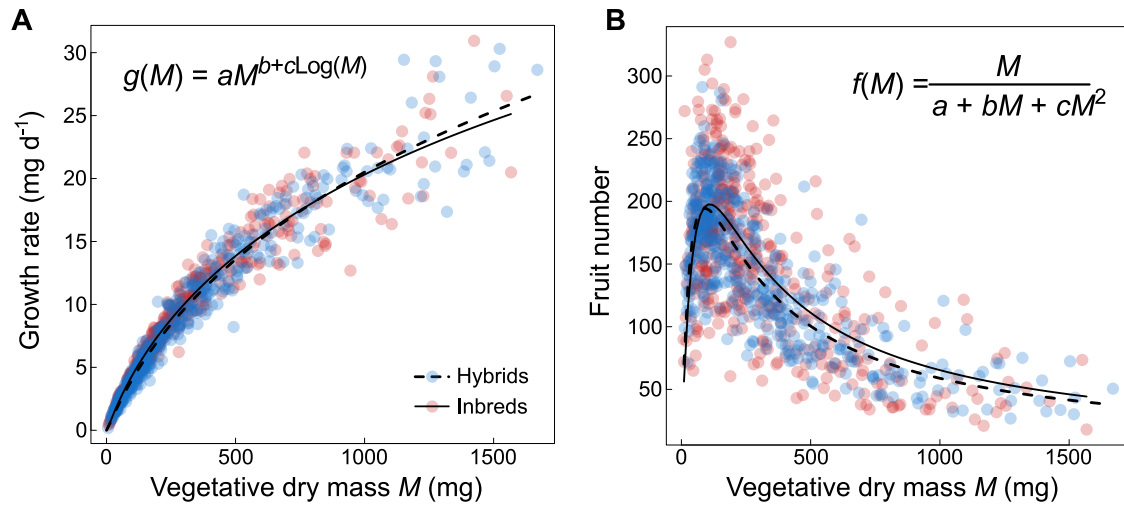


Figure 5.

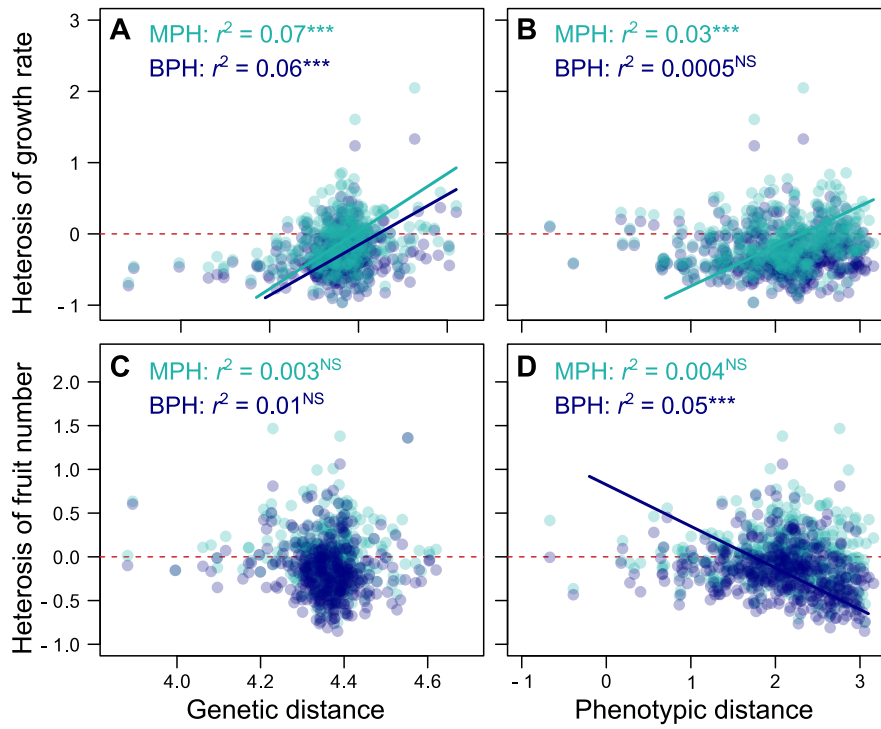


Figure 6.

