# Detecting directional epistasis and dominance from cross-line analyses in alpine populations of *Arabidopsis thaliana*

**Running title:** Directionality of non-additive genetic effects

Arnaud Le Rouzic [1*], Marie Roumet [2,3], Alex Widmer [3], Josselin Clo [4*]

[1] Université Paris-Saclay, CNRS, IRD, UMR Evolution, Génomes, Comportement et Ecologie, France. ORCID 0000-0002-2158-3458, arnaud.le-rouzic@universite-paris-saclay.fr

[2] CTU Bern, University of Bern, Bern, Switzerland. ORCID 0000-0001-6655-1428, marie.roumet@ctu.unibe.ch.

[3] Institute of Integrative Biology, ETH Zurich, Zürich, Switzerland. ORCID 0000-0001-8253-5137, alex.widmer@usys.ethz.ch.

[4] Univ. Lille, CNRS, UMR 8198—Evo-Eco-Paleo, Lille, France. ORCID 0000-0002-3295-9481, josselin.clo@gmail.com

*corresponding authors

**Data availability:** The data and code used to perform the analyses are available at https://github.com/lerouzic/Epicross.

**Authors contributions:** J.C. initiated the project. M.R and A.W. collected the data. A.L.R. developed the statistical framework and J.C. and A.L.R. analyzed the data. All authors wrote the first draft and edited the manuscript.

**Conflict of interest:** The authors have no conflict of interest to declare.

**Abstract:** The contribution of non-additive genetic effects to the genetic architecture of fitness, and to the evolutionary potential of populations, has been a topic of theoretical and empirical interest for a long time. Yet, the empirical study of these effects in natural populations remains scarce, perhaps because measuring dominance and epistasis relies heavily on experimental line crosses. In this study, we explored the contribution of dominance and epistasis in natural alpine populations of *Arabidopsis thaliana*, for two fitness traits, the dry biomass and the estimated number of siliques, measured in a greenhouse. We first found that, on average, crosses between inbred lines of *A. thaliana* led to mid-parent heterosis for the dry biomass, but outbreeding depression for the estimated number of siliques. While heterosis for the dry biomass was due to dominance, we found that outbreeding depression for the estimated number of siliques could be attributed to the breakdown of beneficial epistatic interactions. We simulated and discussed the implication of these results for the adaptive potential of the studied populations, as well as the use of line-cross analyses to detect non-additive genetic effects.

**Keywords:** Non-additive genetic effects, dominance, epistasis, quantitative genetics, self-fertilization, *Arabidopsis*.

## INTRODUCTION

The evolutionary forces underlying divergence between populations, in particular the relative strength of genetic drift compared to natural selection can be revealed by studying the consequences of hybridization (Lynch, 1991; Demuth and Wade, 2005; Fitzpatrick, 2008). For example, when the evolution of populations under study has been mainly driven by genetic drift (e.g. due to demographic events or mating systems, Barrett *et al.*, 2014), populations have likely fixed deleterious mutations (Kimura *et al.*, 1963). Hybrids coming from hybridization events between populations that have accumulated (partially) recessive deleterious mutations are expected to display heterosis, *i.e.* an increase in fitness in $F_1$ hybrids compared to the average fitness of their parents (Crow, 1948; Lynch, 1991; Glémin, 2003). When selection is stronger than drift, divergent populations are expected to accumulate genetic incompatibilities and hybrids may perform poorly compared to the parental populations, in $F_1$ and/or subsequent generations (in general the $F_2$), a phenomenon called outbreeding depression (Lynch, 1991). Several non-exclusive genetic architectures of fitness can lead to outbreeding depression: chromosomal rearrangement, which leads to the production of aneuploid gametes in heterozygotes (Lande, 1985; Charlesworth, 1992; Kirkpatrick and Barton, 2006); under-dominance, which leads to lower fitness for heterozygotes compared to the homozygote genotypes (Schierup and Christiansen, 1996); and negative epistatic interactions among divergent alleles that are brought together for the first time in hybrids (Dobzhansky, 1937; Lynch, 1991; Demuth and Wade, 2005).

Estimating non-additive genetic effects is also important because they are expected to modify how a population responds to selection (Kelly, 1999; Carter *et al.*, 2005). Epistasis has the potential to modify the short- and long-term adaptive potential of a species (Cheverud &

Routman, 1995; Carter *et al.*, 2005; Hansen, 2015). Positive epistasis (i.e. epistatic interactions that increase traits' values) tends to amplify the genetic variance of the trait, and consequently, the capacity of the population to respond to selection will also increase (Carter *et al.*, 2005). Negative epistasis (i.e. epistatic interactions that decrease traits' values) will have the opposite effect. Despite its potential effect on response to selection, estimates of epistasis for morphological traits are currently rare and not consistent (Le Rouzic, 2014). Pavlicev *et al.* (2010) found negative epistasis for body-composition traits in *Mus musculus*, while in several plant species, epistasis tended to be positive, on average, for different floral morphological traits (Johansen-Morris & Latta, 2006; Monnahan & Kelly, 2015; Oakley *et al.*, 2015; Clo *et al.*, 2021). Epistasis on fitness and fitness-related traits has received substantially more interest. Theoretical work has indeed repeatedly pointed out that epistasis on fitness should drive many diversity-generating mechanisms, including the evolution of sex, recombination, and mutation rates (Phillips *et al.*, 2000). Yet, empirical estimates don't show a clear directionality (Maisnier-Patin *et al.*, 2005; Kouyos *et al.*, 2007; Bakerlee *et al.*, 2022).

Dominance in quantitative genetics received less theoretical attention because its consequences on the evolutionary potential of a species are complex (Walsh & Lynch, 2018). Dominance can modify the adaptive potential of a trait (Clo *et al*., 2019, Clo & Opedal 2021, Sztepanacz *et al*., 2023). However, its effect depends on inbreeding (Kelly, 1999): dominance can contribute to the covariance between parents and offspring in inbred populations, which does not occur under random mating (Falconer, 1996; Lynch & Walsh, 1998). The adaptive potential of inbred populations is thus not only described by the additive variance (Cockerham & Weir, 1984; Wright & Cockerham, 1985), as dominance contributes to the evolvability of a trait (Clo & Opedal, 2021). Dominance is routinely observed in fitness traits, as shown by the ubiquity of inbreeding depression (Charlesworth & Willis, 2009), but less is known about morphological traits, for which dominance can be either beneficial or deleterious, depending

on studies and traits measured (Shaw *et al.*, 1998; Kelly & Arathi, 2003; Oakley *et al.*, 2015; Clo *et al.*, 2021).

Several experimental protocols have been proposed to detect these non-additive genetic effects (see Le Rouzic, 2014 for a list of methods to detect epistasis). Among all, line-cross analyses have been the most used (Demuth & Wade, 2005), probably due to the simplicity of the crossing design and associated statistics. This method requires the plants to be grown and measured in controlled conditions (generally greenhouses) to minimize the environmental variance and maximize the statistical power to detect genetic effects (Walsh & Lynch, 2018).

Predominantly selfing species are of particular interest when studying the phenotypic and fitness consequences of hybridization. By organizing populations in a mosaic of repeated fully homozygous genetic lines (Siol *et al.*, 2008; Volis *et al.*, 2010; Jullien *et al.*, 2019), self-fertilization simplifies the dissection of the genetic architecture of hybrid fitness. Indeed, hybridization between inbred lines produces $F_1$ hybrids that are expected to be fully heterozygous at all sites where the parents differ. At each subsequent generation of selfing, the heterozygosity is divided by two, allowing one to decipher the relative contribution of dominance and epistasis to hybrids' performance (Demuth and Wade, 2005; Fitzpatrick, 2008).

*Arabidopsis thaliana* is a natural choice to study the genetic architecture of fitness and the consequences of non-additive genetic effects in plants. *A. thaliana* (L.) Heyhn. (Brassicaceae) is native to Eurasia and North Africa but is now widely distributed throughout the Northern hemisphere (Hoffmann, 2002). This species occurs in many mountain ranges, and has been reported along a wide altitudinal range, from sea level up to 2000 m in the central Alps (Hoffmann, 2002). Unlike close relatives, *A. thaliana* is a predominantly self-fertilizing, annual species; average outcrossing rates in natural populations have been reported to vary between 0.3% and 2.5% (Abbott & Gomes, 1989; Bergelson *et al.*, 1998; Picó *et al.*, 2008).

In this study, we explored the genetic architecture of two fitness traits in natural Alpine populations of *A. thaliana*, the dry biomass and the estimated number of siliques. We first found that, on average, crosses between inbred lines of *A. thaliana* led to heterosis for the dry biomass, but outbreeding depression for the estimated number of siliques. We found that heterosis for the dry biomass could be attributed to a positive effect of dominance. This is likely due to the masking of recessive deleterious mutations segregating in the different inbred lines. For the estimated number of siliques, however, we found that outbreeding depression was likely due epistatic interactions. We simulated the response to selection of our traits and found that both dominance and epistasis modified the response to selection compared to additive scenarios.

## MATERIAL AND METHODS

### Study populations

We studied six natural alpine populations located along an altitudinal gradient in the Alps in the Saas Valley (Valais, Switzerland). Focal populations (Table S1) were selected from those studied by Luo *et al.* (2015). Three populations are from low altitudes (*i.e.* altitudes ranging from 850 to 1000m) and three from close to the high-elevational range margin of the species in the Alps (*i.e.* altitudes ranging from 1792 to 2012m). Distances among populations ranged from 0.8 to 25.8 km, with average distances of 6.3 km among low-altitude and 1.9 km among high-altitude populations. For our crossing experiment, we used offspring of plants collected in 2007 that were propagated in the greenhouse for three generations by single-seed descent to standardize maternal effects.

**Experimental design and measured traits**

Seeds of plants originally collected in 2007 and propagated for three generations by single seed descent in a greenhouse were used as parental lines in the crossing experiment. From each of the six study populations we randomly selected four parental lines with different genotypes based on results reported in Luo *et al.* (2015). Briefly, Luo *et al.* (2015) genotyped individuals using twenty-two microsatellite markers (Table S2); the populations' structure was reported in their Table 1 and Table S1. In the following section, when an analysis is referring to be done at the "lines" scale, it means we compared results between parental lines (seeds propagated for three generations by single seed descents). When we are referring to analyses done at the "populations" scale, we compared results between populations. All parental lines were bred to produce four offspring categories: (i) spontaneous self-fertilization, leading to the same parental genotype, (ii) outcrossing with pollen from another parental line from the same local population, (iii) outcrossing with a parental line from another population of the same altitude, and (iv) outcrossing with a parental line from another population of the other altitude. For each parental line, self-fertilization and the three cross types were realized/ on four different plants. Crosses were performed using the pollen of plants exclusively grown for pollen production. Parental lines were randomly matched without replacement such that all four parental lines of each population were included as both seed and pollen parents in all cross types (Table S3). The F2 generation was produced by the spontaneous self-fertilization of one individual from each F1 family (for an overview of the crossing design see Table S3). Crosses were realized by emasculating and manually pollinating 4 to 6 flowers per female/mother plant. To avoid uncontrolled cross-pollination between neighboring plants, unused flowers were removed, and pots were individually packed in Arabisifter floral sleeves (Lehle Seeds, Round Rock, Texas,

USA). To ensure self-fertilization and limit accidental crossing events, selfing plants were also packed in floral sleeves during the flowering stage. No control emasculations were performed.

Performance and phenotypic variation in all F1 and F2 families were assessed in a single large greenhouse experiment conducted in spring 2014. Seeds were sown on March 4th 2014 and stratified at 4 °C in the dark for six days. Plants were then grown in a greenhouse at ETH Zurich research station Lindau-Eschikon under long-day conditions (*i.e.* $10^4$ Lux light for 16 h, dark for 8 h; 22 °C/18 °C day/night temperatures). From each parental line, we grew 12 parental plants, 24 F1 and 48 F2 offspring. The F1 included six offspring derived from selfing and six offspring derived from each of the three cross types. The F2 generation encompassed six selfed offspring and 14 offspring from each of the three cross types. In total, the experiment encompassed 1728 plants in total (for details see Table S3).

Plants were grown individually in 7×7×8 cm pots randomly arranged in two greenhouse compartments, and were filled with Biouniversalerde (Oekohum GmbH, Herrenhof, Switzerland), an all-purpose soil without peat. Within each greenhouse compartment, pots were randomly arranged in 24-pot-trays. To avoid edge effects, trays were placed on tables next to each other and surrounded by “border plants” (*i.e.* plants derived by self-pollination from the study populations, sown and grown under the same conditions as the experimental plants). Trays with experimental plants were randomized twice a week until the maturation of siliques. All plants were harvested on July 1st, 2014, approximately four months after germination, when all plants that reached the flowering stage started to dry. Plants were first dried for 48h at 45°C. We then measured the dry biomass and estimated the number of siliques per plant. To estimate the number of siliques, we first separated the different branches and isolated the reproductive sections of all branches (*i.e.* the parts of the branches carrying fruits); second, we weighted the reproductive ('reproductive weight') sections of all branches of each individual

together to the nearest 0.0001g using a Mettler AE 240 analytical balance. Third, we assessed the estimated number of siliques along three randomly selected and weighted reproductive sections and estimated the number of siliques per gram (‘silique density'); fourth, we estimated the total number of siliques produced per plant ('silique number') as the product of the 'silique density' and the 'reproductive weight'. This last measurement, the 'silique number', was used as a proxy for individual fecundity.

**Genetic model**

Traditional line cross models consider only two parental lines and generally define genetic effects by measuring the difference between parental lines and F1 to the F2 generations. Using the F2 generations as a reference generally simplifies the mathematical expressions (Lynch & Walsh, 1998). As we aim here to analyze several line crosses at once, we reparametrized this model by taking the grand mean of the parents μ as a reference (Table S4, Figure S1). The average phenotypic means of $P_i$ and $P_j$, as well as their intercrosses $F_{1,\ ij}$ and $F_{2,\ ij}$ can be expressed as:

$$P_i = \mu + A_i$$

$$P_j = \mu + A_j$$

$$F_{1,\ ij} = \mu + ½\ A_i + ½\ A_j + 2D_{ij} – AA_{ij}$$

$$F_{2,\ ij} = \mu + ½\ A_i + ½\ A_j + D_{ij} – AA_{ij}$$

This setting defines one additive effect A per parent, and as many dominance (D) and additive-by-additive (AA) epistasis parameters as independent crosses. In the absence of backcrosses, additive-by-dominance epistatic effects cannot be identified and are merged with additive effects. Dominance-by-dominance interactions, as well as higher-order epistatic terms, had to be ignored. Note that positive dominance tends to generate heterosis (F1 and F2 generations larger than the mid-parent), while positive epistasis tends to generate outbreeding depression (F1 and F2 lower than the mid-parent).

Line-cross models aim to measure individual deviations from additivity, and were thus analyzed with a fixed-effect linear model:

$$z_k = \mu + \tfrac{1}{2} A_{P(k)} + \tfrac{1}{2} A_{Q(k)} + d_k D_{P(k)Q(k)} + aa_k AA_{P(k)Q(k)} + e_k$$

for individual k of phenotype $z_k$, of parents from parents P(k) and Q(k), with $d_k = aa_k = 0$ if $P(k) = Q(k)$, $aa_k = -1$ if $P(k) \neq Q(k)$ (k is from an intercross $F_1$ or $F_2$), and $d_k = 2$ (or =1) if k results from an $F_1$ (or an $F_2$) intercross. $e_k$ is a Gaussian-distributed residual of variance $V_e$.

Four models of various complexity were fit to each dataset: Additive (only the additive terms $A_i$ were considered), Dominance ($A_i$ and $D_{ij}$), Epistasis ($A_i$ and $AA_{ij}$), and Full model ($A_i$, $D_{ij}$, and $AA_{ij}$). The four models were compared by a model selection procedure based on the Akaike

Information Criterion (Anderson & Burnham, 2004) of the best fit model; AIC differences larger than 2 units were considered to be a significantly poorer fit to the data. Models were fit independently on dry mass and estimated number of siliques, and both genetic differentiation levels (Lines and Population) were considered.

**Simulated adaptive trajectories**

We used the best model and the associated non-additive genetic effects to simulate the consequences of non-additivity and inbreeding on the response to selection for the traits of interest, compared to an additive model. To do so, we used a two-locus and two-alleles model. We assumed infinite populations, linkage equilibrium, and genotype frequencies were always at equilibrium (accounting for selfing, i.e. the frequency of heterozygotes was p(1-p)(1-s)/(2-s) where s is the selfing rate, Hartl & Clark 1989). Fitness was proportional to the phenotype; the lowest fitness genotype was 80% that of the best. Genotype-phenotype maps were parameterized according to the most supported model (additive effects and dominance for weight, additive and AA epistasis for siliques), and parameterized according to a traditional F2 model:

| | Locus 1 | | |
|---|---|---|---|
| Locus 2 | A1A1 | A1B1 | B1B1 |
| A2A2 | μ - a - d + aa | μ - a/2 | μ - aa + d |
| A2B2 | μ - a/2 | μ + d | μ + a/2 |
| B2B2 | μ - aa - d | μ + a/2 | μ + a - d + aa |

When non-zero, parameter values were the average of pairwise effects (red bars in Fig 2): for weight, μ = 0.43g, a=0.11g, d=0.03g; for siliques: μ=779, a=262, aa=83. In this setting, μ stands for the mean random-mating F2 population. Starting allele frequencies were ½ at both loci; the phenotype of starting populations was not necessarily at μ because of selfing, which induces non-Hardy-Weinberg frequencies.

## RESULTS & DISCUSSION

We found that performing the analyses at the scale of the genetic lines or at the scale of the populations gave compatible results, although population-level analysis had lower statistical power due to the smaller sample size. As within-population variation was often of the same magnitude as between-population variation, averaging lineages within a population was not justified, and we decided to present the results for the genetic lines in the main text, the results at the population scale being available as supplementary material (Table S4, Figures S2 and S3).

### *Consequences of hybridization on dry biomass and the production of siliques*

Our first question was related to the consequences of non-additive genetic effects on the phenotype of hybrid (F1 and F2) populations compared to their parents. The raw phenotypes of parental lines and within-population crosses are available as supplementary materials (Figure S3). We found that, on average, $F_1$ hybrids exhibited heterosis for the dry

biomass (Figure 1), with an increase of 9.6% (0.454g in $F_1$ hybrids), compared to the mean parental value of 0.414g. This amount of heterosis is in line with what is found in other populations of *A. thaliana* (Oakley *et al*., 2015), and in other predominantly selfing species (Rhode & Cruzan, 2005; Dolgin *et al.*, 2007; Volis *et al.*, 2011; Gimond *et al.*, 2013; Clo *et al.*, 2021), for different fitness proxies (dry biomass, fruits and seeds production). In contrast, $F_2$ hybrids had similar mean dry biomass to the parents (0.403g, -2%).

On the other hand, $F_1$ and $F_2$ hybrids exhibited outbreeding depression for the estimated number of siliques (Figure 1), with respectively a decrease of 5.5% and 11.3% in $F_1$ and $F_2$ hybrids, compared to the mean parental value. This is slightly lower than other values found for different fitness proxies (seeds and fruits production, germination rate etc.) in other predominantly selfing species (Rhode & Cruzan, 2005; Dolgin *et al.*, 2007; Volis *et al.*, 2011; Gimond *et al.*, 2013; Oakley *et al.*, 2015; Clo *et al.*, 2021; Soto *et al*., 2023).

These opposite patterns for dry biomass and fruit number could be considered a surprising result since both traits are often considered fitness proxies, and are generally positively correlated (see Younginger *et al.*, 2017 for a review). However, such an observation is not unheard of. Studies in natural and laboratory accessions of *Arabidopsis thaliana* also found heterosis patterns for the dry mass and outbreeding depression for a fitness proxy (pollen viability in Nasrallah *et al.*, 2000; seed production in Barth *et al.*, 2003; fruit production Vasseur *et al.*, 2019). In the sister species *A. lyrata*, Li *et al.* (2019) also found that selfing populations exhibit heterosis for above- and below-ground biomass, and a slight outbreeding depression pattern for fitness (measured as the probability of bolting) in outcrossed progeny of selfing populations. Finally, Clo *et al*. (2021) found that in the predominantly selfing species *Medicago truncatula*, hybridization between inbred lines leads to heterosis for dry mass but outbreeding depression for seed production. It is known that environmental factors, such as

nutrients or temperature, have a key role in the transition from vegetative to reproductive stages, like in flowering probability (see Cho et al. 2017 for a review). It is thus possible that the ecological conditions found in the greenhouse, which are more favorable than those in the field, might have modified trade-offs between vegetative growth and investment in reproduction; extrapolating our greenhouse results to natural populations thus relies on the /assumption that environmental conditions do not affect the relative performance of genotypes (limited G⨯E interactions).

We also found that, on average, plants from low-altitude populations or hybrids generated with lines from low-altitude populations were generally heavier and produced more siliques compared to high-altitude plants (Figure S4). However, this result is mainly due to a few inbred lines within each high-altitude population that performed badly, while others were in lines with what is found in low-altitude populations (Figure S3). This altitude effect can be explained by the difference in ecological conditions between low- and high-altitude populations, with for example mean annual temperature and precipitation being very different (see Luo *et al*., 2015 for details).

**Non-additive effects in natural populations of plants**

We found that non-additive effects contribute to the genetic architecture of both traits. For dry mass, we found that the best model explaining the data was the one including additive and dominant genetic effects (Table 1), and the observed pattern of heterosis was mostly due to positive dominant effect (d=+0.03g on average, to be compared with an average additive effect of a=0.11g, Figure 2). Oakley *et al.* (2015) found similar results in crosses between south

European and Scandinavian lineages of *A. thaliana*. The positive dominance likely reflects the positive effects of masking deleterious mutations fixed at different loci in the different selfing lines (Charlesworth 2018). The contribution of epistasis to dry biomass cannot totally be ruled out when the data are analyzed at the population level (Table S4). It indicates that both the masking of deleterious mutations and a potential synergistic effect of masking at different sites could explain the heterosis pattern (Oakley *et al*., 2015; Charlesworth 2018).

For the estimated number of siliques, we found that the best model explaining the data was the one including additive and additive-by-additive epistatic genetic effects (Table 1), and the observed pattern of outbreeding depression was due to positive additive-by-additive epistatic interactions (average aa = 83 siliques, to be compared to an average additive effect a = 262 siliques, Figure 2). The outbreeding depression can be explained by the breakdown of positive additive-by-additive epistatic interactions found in the parental selfing lines during the hybridization events, as found in other species (Rhode & Cruzan, 2005; Johansen-Morris & Latta, 2006; Monnahan & Kelly, 2015; Oakley *et al.*, 2015; Clo *et al.*, 2021).

**Implications for the adaptive potential of Alpine populations of *A. thaliana***

The distinct genetic architecture among the two fitness traits studied here implies different effects on the capacity to respond to selection. By using a two-locus and two-alleles model that mimics the genetic architecture of our traits, we explored the consequences of non-additive effects on the capacity to respond to a hypothetical selection pressure (Figure 3). When the departure to additivity is attributed to dominance,  as for the dry mass, the response to

selection depends on the selfing rate (Figure 3(a)). The fact that selfing populations respond faster to selection is expected, as selfing increases the heritable variance by a factor of (1+F) due to the higher proportion of homozygotes genotypes (Wright, 1921). The fact that dominance *per se* did not affect the response to selection in selfing populations is also expected, because dominance is only expressed in heterozygote genotypes, which are rare in predominantly selfing populations/species such as *A. thaliana.* With inbreeding and dominance, the estimation of adaptative potential is different than in outcrossing species, because the whole genetic variance (in the different components of the genetic diversity) is a better predictor of the capacity to respond to selection than the additive variance (see for example Wright & Cockerham, 1985, Kelly 1999 or Clo & Opedal 2021 for further details). For the estimated number of siliques, we found positive additive-by-additive epistasis. In the short-term, positive epistasis tended to increase the capacity to respond to selection (Figure 3(b)), as the silique production increased faster with epistasis (Figure 3(b)). This is due to the fact that positive epistasis increases the amount of additive variance of a quantitative trait, and, as a result, the capacity to respond to selection (Carter *et al.*, 2005; Monnahan & Kelly, 2015). In the long term, epistasis slightly increased the seed production compared to an additive model (Figure 3(b)), the difference being small because the epistatic coefficient is small compared to additive effects.

**Limits of the statistical model to detect the nature of non-additive effects**

The line cross model we proposed makes it possible to test for the presence of dominance and epistasis, but it has potential drawbacks. First, the fact that plants have to be grown in controlled conditions makes the different traits to be very different from what is found

in natural populations (see for example Weng *et al*., 2021 for estimates in natural populations of *A. thaliana*). This overestimation of the effects may make our predictive response to selection larger than expected in nature, but the pattern should be qualitatively the same. Second, it can estimate epistasis, which is essential for understanding the capacity of populations to respond to selection, but it was not possible to dissect the different forms of epistatic variance (additive-by-additive, additive-by-dominant, and dominant-by-dominant). For inferring all these parameters from cross-line analyses, one needs more crosses than just the $F_1$ and $F_2$ individuals (see Lynch & Walsh, 1998, including reciprocal back-crosses for example, and see Oakley *et al.*, 2015 for a case study). This means that the positive dominance we detected for the dry biomass could be a mixture of negative dominance and epistasis, or just complex forms of epistasis (beyond additive-by-additive interaction effects). As a consequence of this identifiability problem, genetic effect estimates for dominance and epistasis tend to be statistically correlated; in other words, the model has a substantial power to detect departure from additivity, but less power to disentangle dominance and epistasis.

## CONCLUSIONS

Our study highlights the contribution of non-additive genetic effects to the genetic architecture of fitness components. Here, we found that both dominance and epistasis affect the genetic architecture of dry biomass and silique production, leading to heterosis for the dry mass and outbreeding depression for the estimated number of siliques in $F_1$ and $F_2$ hybrids. We however found that non-additive genetic effects remain quantitatively small compared to additive components. Simulations reflecting the genetic architecture of the studied traits, as well as the mating systems of our model species *A. thaliana*, showed that both the non-additive

genetic effect and the selfing rate have a significant influence on the potential to respond to selection. Testing such theoretical predictions with artificial selection experiments is challenging (see Monnahan and Kelly, 2015; for a case study), but it appears as a necessary step to verify whether non-additive theoretical models of adaptation are operational in practice.

## BIBLIOGRAPHY


Abbott, R.J. & Gomes, M.F. 1989. Population genetic structure and outcrossing rate of Arabidopsis thaliana (L.) Heynh. *Heredity* **62**: 411–418.

Anderson, D. & Burnham, K. 2004. Model selection and multi-model inference. *Second. NY: Springer-Verlag* **63**: 10.

Bakerlee, C.W., Nguyen Ba, A.N., Shulgina, Y., Rojas Echenique, J.I. & Desai, M.M. 2022. Idiosyncratic epistasis leads to global fitness–correlated trends. *Science* **376**: 630–635.

Barrett, S. C., Arunkumar, R., & Wright, S. I. 2014. The demography and population genomics of evolutionary transitions to self-fertilization in plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**: 20130344.

Barth, S., Busimi, A.K., Utz, H.F. & Melchinger, A.E. 2003. Heterosis for biomass yield and related traits in five hybrids of Arabidopsis thaliana L. Heynh. *Heredity* **91**: 36–42.

Bergelson, J., Stahl, E., Dudek, S. & Kreitman, M. 1998. Genetic variation within and among populations of Arabidopsis thaliana. *Genetics* **148**: 1311–1323.

Carter, A.J., Hermisson, J. & Hansen, T.F. 2005. The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theoretical population biology* **68**: 179–196.

Charlesworth, D. & Willis, J.H. 2009. The genetics of inbreeding depression. *Nature reviews genetics* **10**: 783–796.

Charlesworth, B. 1992. Evolutionary rates in partially self-fertilizing species. *The American Naturalist* **140:** 126-148.

Charlesworth, B. (2018). Mutational load, inbreeding depression and heterosis in subdivided populations. *Molecular ecology*, **27**(24), 4991-5003.

Cheverud, J.M. & Routman, E.J. 1995. Epistasis and its contribution to genetic variance components. *Genetics* **139**: 1455–1461.

Cho, L.H., Yoon, J. & An, G. 2017. The control of flowering time by environmental factors. *The Plant Journal*, **90**(4), 708-719.

Clo, J., Gay, L. & Ronfort, J. 2019. How does selfing affect the genetic variance of quantitative traits? An updated meta-analysis on empirical results in angiosperm species. *Evolution* **73**: 1578–1590.

Clo, J. & Opedal, Ø.H. 2021. Genetics of quantitative traits with dominance under stabilizing and directional selection in partially selfing species. *Evolution* **75**: 1920–1935.

Clo, J., Ronfort, J. & Gay, L. 2021. Fitness consequences of hybridization in a predominantly selfing species: insights into the role of dominance and epistatic incompatibilities. *Heredity* **127**: 393–400.

Cockerham, C.C. & Weir, B.S. 1984. Covariances of relatives stemming from a population undergoing mixed self and random mating. *Biometrics* 157–164.

Crow, J. F. 1948. Alternative hypotheses of hybrid vigor. *Genetics* **33:** 477.

Demuth, J. P., & Wade, M. J. 2005. On the theoretical and empirical framework for studying genetic interactions within and among species. *The American Naturalist* **165:** 524-536.

Dobzhansky, T. 1982. *Genetics and the Origin of Species* (No. 11). Columbia university press.

Dolgin, E.S., Charlesworth, B., Baird, S.E. & Cutter, A.D. 2007. Inbreeding and outbreeding depression in Caenorhabditis nematodes. *Evolution: International Journal of Organic Evolution* **61**: 1339–1352.

Falconer, D.S. & Mackay, T.F.C. 1996. Introduction to quantitative genetics. Essex. *UK: Longman Group*.

Fitzpatrick, B. M. 2008. Hybrid dysfunction: population genetic and quantitative genetic perspectives. *The American Naturalist* **171:** 491-498.

Glémin, S. 2003. How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution* **57:** 2678-2687.

Gimond, C., Jovelin, R., Han, S., Ferrari, C., Cutter, A.D. & Braendle, C. 2013. Outbreeding depression with low genetic variation in selfing Caenorhabditis nematodes. *Evolution* **67**: 3087–3101.

Hansen, T.F. 2015. Measuring gene interactions. *Epistasis* 115–143. Springer.

Hartl, D. L., & Clark, A. G. 1989. Principles of population genetics. Sinauer Assoc. *Inc, Sunderland, Massachusetts*.

Hoffmann, M.H. 2002. Biogeography of Arabidopsis thaliana (l.) heynh.(Brassicaceae). *Journal of Biogeography* **29**: 125–134.

Johansen-Morris, A.D. & Latta, R.G. 2006. Fitness consequences of hybridization between ecotypes of Avena barbata: hybrid breakdown, hybrid vigor, and transgressive segregation. *Evolution* **60**: 1585–1595.

Jullien, M., Navascués, M., Ronfort, J., Loridon, K., & Gay, L. 2019. Structure of multilocus genetic diversity in predominantly selfing populations. *Heredity* **123:** 176-191.

Kelly, J.K. 1999. Response to selection in partially self-fertilizing populations. I. Selection on a single trait. *Evolution* **53**: 336–349.

Kelly, J.K. & Arathi, H.S. 2003. Inbreeding and the genetic variance in floral traits of Mimulus guttatus. *Heredity* **90**: 77–83.

Kimura, M., Maruyama, T., & Crow, J. F. 1963. The mutation load in small populations. *Genetics* **48:** 1303.

Kirkpatrick, M., & Barton, N. 2006. Chromosome inversions, local adaptation and speciation. *Genetics* **173:** 419-434.

Kouyos, R.D., Silander, O.K. & Bonhoeffer, S. 2007. Epistasis between deleterious mutations and the evolution of recombination. *Trends in ecology & evolution* **22**: 308–315.

Lande, R. 1985. The fixation of chromosomal rearrangements in a subdivided population with local extinction and colonization. *Heredity* **54:** 323-332.

Le Rouzic, A. 2014. Estimating directional epistasis. *Frontiers in Genetics* **5**: 198.

Li, Y., van Kleunen, M. & Stift, M. 2019. Sibling competition does not magnify inbreeding depression in North American Arabidopsis lyrata. *Heredity* **123**: 723–732.

Luo, Y., Widmer, A. & Karrenberg, S. 2015. The roles of genetic drift and natural selection in quantitative trait divergence along an altitudinal gradient in Arabidopsis thaliana. *Heredity* **114**: 220–228.

Lynch, M. 1991. The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution* **45 :** 622-629.

Lynch, M. & Walsh, B. 1998. *Genetics and analysis of quantitative traits*. Sinauer Sunderland, MA.

Maisnier-Patin, S., Roth, J.R., Fredriksson, Å., Nyström, T., Berg, O.G. & Andersson, D.I. 2005. Genomic buffering mitigates the effects of deleterious mutations in bacteria. *Nature Genetics* **37**: 1376–1379.

Monnahan, P.J. & Kelly, J.K. 2015. Epistasis is a major determinant of the additive genetic variance in Mimulus guttatus. *PLoS Genet* **11**: e1005201.

Nasrallah, M.E., Yogeeswaran, K., Snyder, S. & Nasrallah, J.B. 2000. Arabidopsis species hybrids in the study of species differences and evolution of amphiploidy in plants. *Plant Physiology* **124**: 1605–1614.

Oakley, C.G., Agren, J. & Schemske, D.W. 2015. Heterosis and outbreeding depression in crosses between natural populations of Arabidopsis thaliana. *Heredity* **115**: 73–82.

Pavlicev, M., Le Rouzic, A., Cheverud, J.M., Wagner, G.P. & Hansen, T.F. 2010. Directionality of epistasis in a murine intercross population. *Genetics* **185**: 1489–1505.

Phillips, P.C., Otto, S.P., Whitlock, M.C., Wolf, J.D., Brodie, E.D.I. & Wade, M.J. 2000. Beyond the average: the evolutionary importance of gene interactions and variability of epistatic effects. *Epistasis and the evolutionary process* 20–38.

Picó, F.X., Méndez-Vigo, B., Martinez-Zapater, J.M. & Alonso-Blanco, C. 2008. Natural genetic variation of Arabidopsis thaliana is geographically structured in the Iberian peninsula. *Genetics* **180**: 1009–1021.

Rhode, J.M. & Cruzan, M.B. 2005. Contributions of heterosis and epistasis to hybrid fitness. *The American Naturalist* **166**: E124–E139.

Schierup, M. H., & Christiansen, F. B. 1996. Inbreeding depression and outbreeding depression in plants. *Heredity* **77:** 461-468.

Shaw, R.G., Byers, D.L. & Shaw, F.H. 1998. Genetic components of variation in Nemophila menziesii undergoing inbreeding: morphology and flowering time. *Genetics* **150**: 1649–1661.

Siol, M., Prosperi, J. M., Bonnin, I., & Ronfort, J. 2008. How multilocus genotypic pattern helps to understand the history of selfing populations: a case study in Medicago truncatula. *Heredity* **100:** 517-525.

Soto, T. Y., Rojas-Gutierrez, J. D., & Oakley, C. G. 2023. Can heterosis and inbreeding depression explain the maintenance of outcrossing in a cleistogamous perennial?. *American Journal of Botany* **110:** e16240.

Sztepanacz, J. L., Clo, J., & Opedal, Ø. H. (2023). Evolvability, Sexual Selection, and Mating Strategies. in *Evolvability, a unifying concept in evolutionary biology?*

Vasseur, F., Fouqueau, L., De Vienne, D., Nidelet, T., Violle, C. & Weigel, D. 2019. Nonlinear phenotypic variation uncovers the emergence of heterosis in Arabidopsis thaliana. *PLoS biology* **17**: e3000214.

Volis, S., Shulgina, I., Zaretsky, M. & Koren, O. 2011. Epistasis in natural populations of a predominantly selfing plant. *Heredity* **106**: 300–309.

Walsh, B. & Lynch, M. 2018. *Evolution and selection of quantitative traits*. Oxford University Press.

Weng, M. L., Ågren, J., Imbert, E., Nottebrock, H., Rutter, M. T., & Fenster, C. B. 2021. Fitness effects of mutation in natural populations of Arabidopsis thaliana reveal a complex influence of local adaptation. *Evolution* **75**: 330-348.

Wright, A.J. & Cockerham, C.C. 1985. Selection with partial selfing. I. Mass selection. *Genetics* **109**: 585–597.

Younginger, B.S., Sirová, D., Cruzan, M.B. & Ballhorn, D.J. 2017. Is biomass a reliable estimate of plant fitness? *Applications in Plant Sciences* **5**: 1600094.

**Table 1.** Summary of the statistical models fitted to data, when analyses are performed at the scale of genetic lines, for the dry biomass and the estimated number of siliques. In the table, "a" stands for additive, "d" for dominance, and "aa" for additive-by-additive epistasis. ΔAIC is the difference in AIC values between the observed and best models, differences of 2 AIC units or more are generally considered as solid statistical support for the best model.

| Trait | Model | Log(likelihood) | d.f. | AIC | ΔAIC |
|---|---|---|---|---|---|
| dry biomass | a | 535.20 | 28 | -1014.39 | 60.84 |
| | a.d | 630.62 | 93 | -1075.24 | 0.00 |
| | a.aa | 625.43 | 91 | -1068.86 | 6.38 |
| | a.d.aa | 683.88 | 151 | -1065.75 | 9.48 |
| | | | | | |
| Estimated number of silliques | a | -10 729.13 | 28 | 21 514.25 | 98.10 |
| | a.d | -10 624.23 | 93 | 21 434.46 | 18.31 |
| | a.aa | -10 617.08 | 91 | 21 416.15 | 0.00 |
| | a.d.aa | -10 574.11 | 151 | 21 450.22 | 34.06 |

**Figure captions:**

**Figure 1.** Graphical representation of the parental, $F_1$ and $F_2$ values, when analyses are performed at the scale of the genetic lines. **(a)** Empty representation indicating the sign of dominance and epistasis as a function of the position of the data points. **(b)** Location of heterosis and outbreeding depression for F1 (colors) and F2 (hatching) populations. **(c)** Distribution of data for the dry biomass. **(d)** Distribution of data for the estimated number of siliques. The ellipse represents the 95% confidence ellipse of the average across all crosses.

**Figure 2.** Distribution of the dominance (**(a)** and **(b)**) and epistatic (**(c)** and **(d)**) genetic effects when analyses are performed at the scale of the genetic lines, for the dry biomass (**(a)** and **(c)**) and the estimated number of siliques (**(b)** and **(d)**). The red lines indicate mean values.

**Figure 3.** Simplified simulations of hypothetical responses to directional selection on dry biomass **(a)** and estimated number of siliques **(b)**. The additive case (hyphenates lines) is contrasted to two-allele two-locus genetic architectures in which the relevant non-additive effects are accounted for (dominance for dry mass, epistasis for estimated number of siliques; effects were averaged out over all pairs of populations). The selfing rate in *A. thaliana* is about 0.95; simulations with random mating populations (gray) are provided for comparison.

**Figure 1**

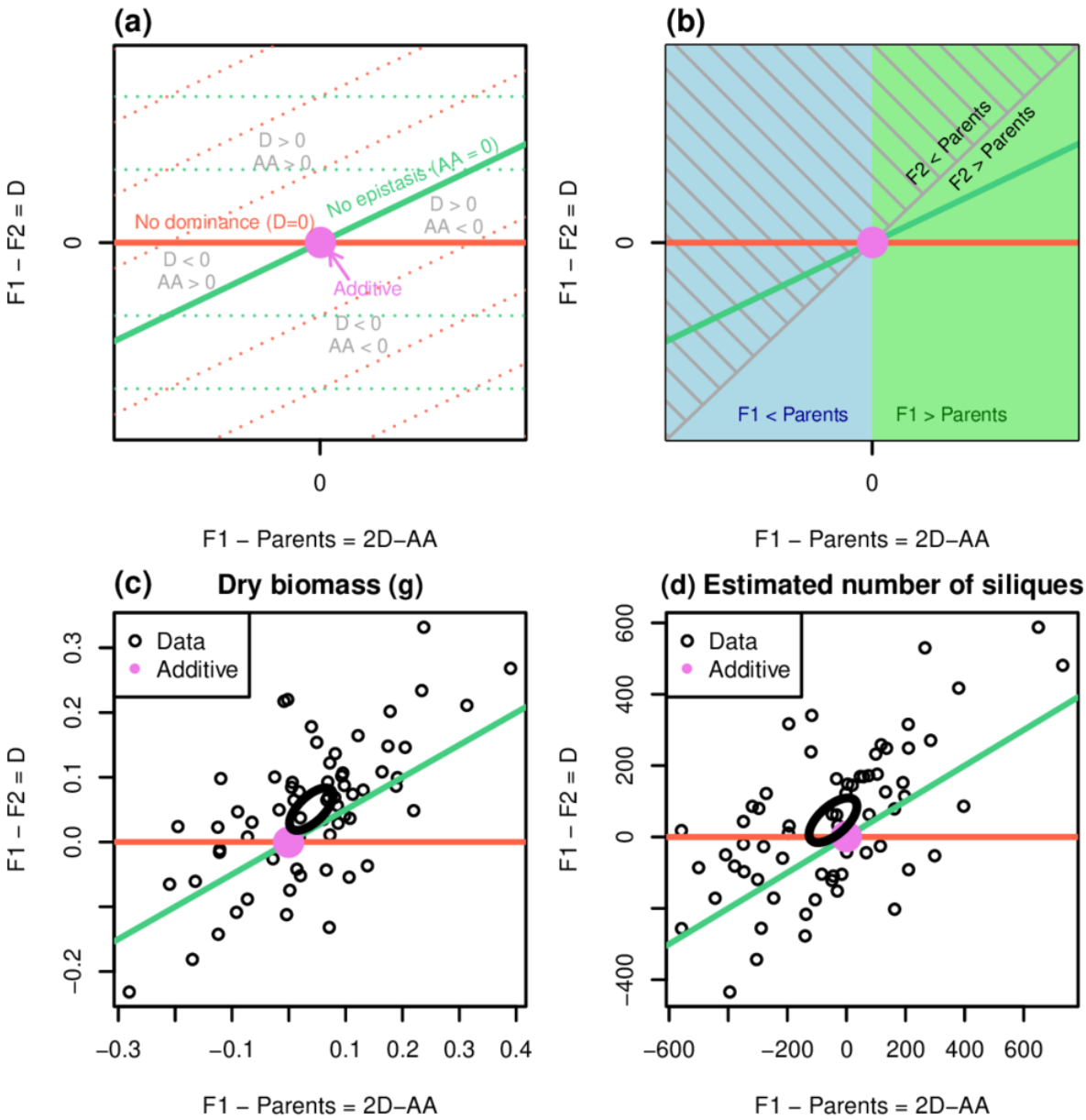
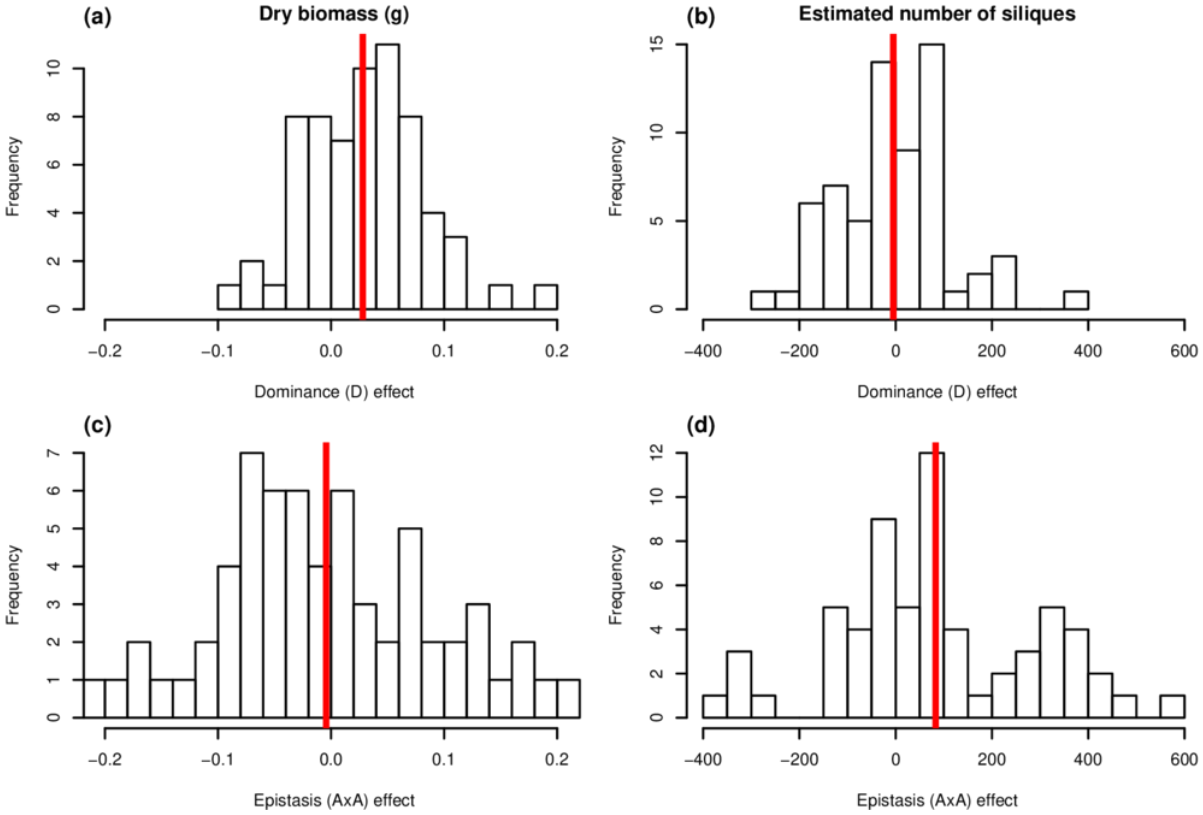
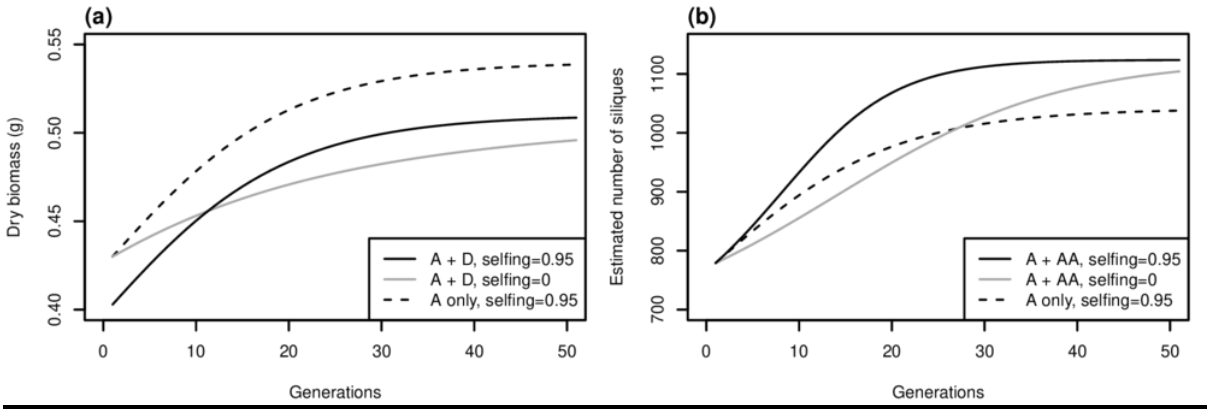

(a)
D > 0
AA > 0
No epistasis (AA = 0)
No dominance (D=0)
D > 0
AA < 0
D < 0
AA > 0
Additive
D < 0
AA < 0
F1 – F2 = D
0
0
F1 – Parents = 2D–AA
(b)
F2 < Parents
F2 > Parents
F1 < Parents
F1 > Parents
F1 – F2 = D
0
0
F1 – Parents = 2D–AA
(c) Dry biomass (g)
Data
Additive
F1 – F2 = D
0.3
0.2
0.1
0.0
–0.2
–0.3
–0.1
0.1
0.2
0.3
0.4
F1 – Parents = 2D–AA
(d) Estimated number of siliques
Data
Additive
F1 – F2 = D
600
400
200
0
–400
–600
–200
0
200
400
600
F1 – Parents = 2D–AA

**Figure 2**


(a)
Dry biomass (g)
Frequency
Dominance (D) effect
(b)
Estimated number of siliques
Frequency
Dominance (D) effect
(c)
Frequency
Epistasis (AxA) effect
(d)
Frequency
Epistasis (AxA) effect

585 **Figure 3**




586