Significance of Linkage Disequilibrium and Epistasis on the Genetic Variances in Non-Inbred and Inbred Populations

José Marcelo Soriano Viana (jmsviana@ufv.br)
Federal University of Viçosa
Antonio Augusto Franco Garcia
University of São Paulo

Research Article

Keywords: linkage disequilibrium, epistasis, inbreeding, genetic variances.

Posted Date: June 29th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-646130/v1

Significance of linkage disequilibrium and epistasis on the genetic variances in non-inbred and inbred populations

José Marcelo Soriano Viana[1*] and Antonio Augusto Franco Garcia[2]

[1]Department of General Biology, Federal University of Viçosa, 36570-900, Viçosa, MG, Brazil.

[2]Department of Genetics, Luiz de Queiroz College of Agriculture, University of São Paulo, 13418-900, Piracicaba, SP, Brazil.

[*]Corresponding author. E-mail: jmsviana@ufv.br. ORCID: https://orcid.org/0000-0002-5063-4648.

Abstract

Background The influence of linkage disequilibrium (LD), epistasis, and inbreeding on the genotypic variance continues to be an important area of investigation in genetics and evolution. Although the current knowledge about biological pathways and gene networks imply that epistasis is important in determining quantitative traits, the empirical evidence for a range of species and traits is that the genetic variance is most additive. This is confirmed by some recent theoretical studies. However, because these investigations have assumed linkage equilibrium, only additive effects, or simplified assumptions for the two- and high-order epistatic effects, the objective of this investigation was to provide additional information about the impact of LD and epistasis on the genetic variances in non-inbred and inbred populations, using a simulated data set.

Results The epistatic variance in generation 0 corresponded to 1 to 10% of the genotypic variance, with 30% of epistatic genes, but it corresponded to 5 to 45% assuming 100% of epistatic genes. After 10 generations of random cross or selfing the ratio epistatic variance/genotypic variance increased in the range of 15 to 1,079%. The epistatic variances are maximized assuming dominant epistasis, duplicate genes with cumulative effects, and non-epistatic gene interaction. A minimization occurs with complementary, recessive, and dominant and recessive epistasis. In non-inbred populations, the genetic covariances have negligible magnitude compared with the genetic variances. In inbred populations, excepting for duplicate epistasis, the sum of the epistatic covariances was in general negative and with magnitude higher than the non-additive variances, especially under 100% of epistatic genes.

Conclusions The LD level for genes, even under a relatively low gene density, has a significant effect on the genetic variances in non-inbred and inbred populations. Assuming digenic epistasis, the additive variance is in general the most important component of the genotypic variance in non-inbred and inbred populations. The ratio epistatic variance/genotypic variance is proportional to the percentage of interacting genes and increases with random cross and selfing. In general, the additive

x additive variance is the most important component of the epistatic variance. The maximization of the epistatic variance depends on the allele frequency, LD level, and epistasis type.

Keywords: linkage disequilibrium, epistasis, inbreeding, genetic variances.

Background

The basic knowledge on the genetics of the quantitative traits was provided by RA Fisher [1], including the partition of the genotypic value in effects due to individual genes, allelic interaction (dominance), and non-allelic interaction (epistasis). Further, he also recognized the significance of the linkage phase between genes on the population variance and on the correlation between relatives. The influence of linkage disequilibrium (LD), epistasis, and inbreeding on the genotypic variance continues to be an important area of investigation in genetics and evolution [2-4]. Assuming linkage equilibrium and three to five loci interaction, A Maki-Tanila and WG Hill [4] concluded that most of the genotypic variance is additive, regardless of order of interaction, allele frequencies, and type and magnitude of interaction effects. Another main finding was that the majority of the epistatic variance is due to digenic interactions. Assuming LD, WG Hill and A Maki-Tanila [3] showed that variances are generally higher with positive LD and that the ratio epistatic variance/genotypic variance is largest with negative LD. Both studies showed that the epistatic variance is increased by increasing the heterozygosity. However, this has no impact on the relative magnitude of the epistatic variance because the additive and epistatic variances increase in similar proportions.

Based on the additive model, J Clo, J Ronfort and D Abu Awad [2] showed that assuming stabilizing selection and high mutation rates, self-pollinated populations are able to accumulate genetic variation through negative LD. Using a meta-analysis of quantitative traits heritability, J Clo, L Gay and J Ronfort [5] confirmed previous theoretical and empirical evidences that self-pollinated populations exhibit lower levels of additive variance for quantitative traits. However, the decrease in the additive variance is compensated by the non-additive components of the genotypic variance. Because of negative consequences (inbreeding depression), geneticists agree that inbreeding should be efficiently controlled to maintain adequate genetic diversity in the populations [6, 7]. However, self-pollination has been deliberately used in maize hybrid breeding (currently to a lesser extent due to the doubled-haploid technology). For self-pollinated crops, the development of

varieties involves selection over generations of increasing inbreeding. In these populations the inbreeding has an impact on the genetic variances and covariance between relatives [8].

Although the current knowledge about biological pathways and gene networks imply that epistasis is important in determining quantitative traits, the empirical evidence for a range of species and traits is that the genetic variance is most additive [9, 10]. Based on theoretical models, WG Hill, ME Goddard and PM Visscher [10] concluded that this occurs because high difference of allelic frequencies. They also concluded that, in outbred populations, the detection of epistasis is difficult unless the epistatic effects are large and the gene frequencies are intermediate. TFC Mackay [9] emphasizes that, because epistasis regularly determine quantitative traits, it has consequences for plant and animal breeding, evolutionary biology, and human genetics. Recent studies on genomic selection and GWAS including epistasis have confirmed that most of the genetic variance is additive [11-14]. However, incomplete LD at low marker density can indicate epistasis when the trait determination is purely additive [15].

The most important quantitative genetics theory for modelling epistasis was developed by O Kempthorne [16]. CC Cockerham [17] also provided a significant contribution. If modelling only inbreeding, LD, or epistasis is a difficult task for the quantitative geneticists, jointly modelling the three events is a challenge. An impressive approach for two genes theory in quantitative genetics assuming inbreeding, LD, and epistasis was presented by BS Weir and CC Cockerham [18]. Because of the complexity of the expressions for the genetic variances and covariance between relatives, they concluded that “the result is of little use”. That is, the functions do not allow assessing the influence of LD, epistasis, and inbreeding on the genetic variability and the degree of relationship in the populations. Further, because recent investigations based on theoretical models have assumed linkage equilibrium, only additive effects, or simplified assumptions for the two- and high-order epistatic effects, the objective of this study was to provide additional information about the impact of LD and epistasis on the genetic variances in non-inbred and inbred populations, using a simulated data set.

Results

The analysis of the parametric LD in the populations shows that the LD level depends mainly on the gene density (Additional File Figure 1). The higher LD level was observed under high gene density (one gene each cM). Regardless of the gene density, the LD level is generally higher for the closest genes. Because the LD is predominantly positive, 10 generations of random cross significantly decreased the LD level of the populations. The decrease was higher for the density of one gene each five cM, regardless of the population (approximately 95% for r^2, on average). The average r^2 decrease for the density of one gene each cM was 81%. The LD level showed only a slight decrease after 10 generations of selfing, regardless of the population (approximately 14% for r^2, on average).

To characterize the magnitude of the genotypic variance components in non-inbred and inbred populations with contrasting LD levels, under no epistasis, we assumed a density of one gene each five cM. In generation 0, the r^2 in the high LD population is 2,395 times greater than the LD in the low LD population, on average. Compared with the populations with intermediate LD, the r^2 in the high LD population, generation 0, is 372 and 502 times greater, on average. Because the populations with high and low LD levels have an average allele frequency of 0.5, the decrease in the population mean due to inbreeding and the genotypic and additive variances are maximized, relative to the populations with average allele frequency lower (0,3) or higher (0,7) than 0.5. The same is true for the dominance variance in the non-inbred populations. After 10 generations of selfing, the decreases in the population means were 15 and 17% for the populations with low and high LD level, respectively (Additional File Figure 2). Regardless of the LD level and the degree of inbreeding, the additive variance is the most important component of the genotypic variance. The significance of the LD level is impressive on the additive and dominance variances. The additive variance in the population with high LD is 6.8 times greater than the additive variance in the population with low LD in generation 0, 2.9 times greater after 10 generations of random cross, and 5.7 times greater after 10 generations of selfing. Concerning the dominance variance and the

covariance between additive and dominance values, there is a lower difference between their magnitudes in the populations with low and high LD levels. In the non-inbred populations, the dominance variance assuming high LD is approximately two times greater than the dominance variance under low LD, regardless of the generation. In the populations with intermediate LD level, the decreases in the population mean due to inbreeding are similar. In both improved and not improved populations, there was also a significant decrease in the additive variance with random crosses (approximately 60%) and an increase with selfing (approximately 60% too). The additive variance is greater in the not improved population, regardless of the generation. In both populations the additive variance is in general intermediate to the values observed for the populations with high and low LD level. The dominance variance significantly decreased with random cross or selfing, regardless of the level of LD (approximately 12 to 97%).

To characterize the components of the genotypic variance in non-inbred and inbred populations with high LD level, under epistasis, we also assumed the density of one gene each five cM. Regardless of the type of epistasis and the percentage of interacting genes, there are non-significant changes in the population mean along 10 generations of random cross (–0.5 to 0.3%; remember that the average decrease in the r^2 values was approximately 95%) (Additional File Figure 3). With 10 generations of selfing, regardless of the percentage of epistatic genes, except for duplicate and dominant epistasis with 100% of interacting genes, the inbreeding decreased the population mean in 2 to 28% (remember that the decrease assuming no epistasis was 17%).

Regardless of the type of epistasis, the ratio epistatic variance/genotypic variance is proportional to the percentage of the epistatic genes. The epistatic variance in generation 0 corresponded to 1 to 10% (dominant epistasis) of the genotypic variance, with 30% of epistatic genes, but it corresponded to 5 to 45% (duplicate epistasis) assuming 100% of epistatic genes (Additional File Figures 4 to 10). In general, irrespective of the type of epistasis and the percentage of epistatic genes, after 10 generations of random cross or selfing the ratio epistatic variance/genotypic variance increased in the range of 15 to 1,079%. This occurred because the

decrease in the genotypic variance was much higher than the decrease in the epistatic variance with random cross. With selfing, this occurred because the increase in the genotypic variance was much lower than the increase in the epistatic variance or because the genotypic variance decreased while the epistatic variance increased. With one exception, regardless of the type of epistasis and the percentage of epistatic genes, the most important component of the genotypic variance is also the additive variance. The additive variance decreased with random cross and increased with selfing. With duplicate epistasis and 100% of epistatic genes, the additive x additive variance was higher than the additive variance, after three generations of selfing. Except for dominant epistasis, duplicate genes with cumulative effects, and non-epistatic genic interaction, the additive variance was 1.1 to 6 times greater assuming 30% of epistatic genes, compared with 100% of epistatic genes, for both random cross and selfing. Assuming dominant epistasis, duplicate genes with cumulative effects, and non-epistatic genic interaction, the additive variance was 1.5 to 2.7 times higher with 100% of interacting genes, compared with 30% of interacting genes.

For the epistatic variances, their magnitudes are much lower than the additive variance (Additional File Figures 4 to 10). The additive x additive variance is the most important epistatic variance. Generally, an insignificant variation in the epistatic variances was observed throughout 10 generations of random cross (−13 to 6%), regardless of the type of epistasis and the percentage of the epistatic genes. A significant increase in the additive x additive, additive x dominant, and dominant x additive variances occurred with selfing (114 to 863%), regardless of the percentage of epistatic genes and the type of epistasis. When inbreeding increased, the dominant x dominant variance significantly decreased in the population with high LD (76 to 86%) but increased in the other populations (11 to 175%). The epistatic variances are maximized assuming dominant epistasis, duplicate genes with cumulative effects, and non-epistatic gene interaction. A minimization of the epistatic variances occurs with complementary, recessive, and dominant and recessive epistasis. In non-inbred populations, the genetic covariances have negligible magnitude compared with the genetic variances. In inbred populations, excepting for duplicate epistasis, the

sum of the epistatic covariances was in general negative and with magnitude higher than the non-additive variances, especially under 100% of epistatic genes.

For the populations with intermediate and low LD levels, the previous inferences holds but the genotypic and genetic variances are generally lower than the values for the population of high LD level, regardless of generation, type of epistasis, and percentage of epistatic genes, as exemplified assuming 30% of epistatic genes showing all types of epistasis (Figure 1). With no exception, the additive variance is also the most important component of the genotypic variance, regardless of the generation. Further, assuming an admixture of the epistasis types and 30% of interacting genes, the ratio epistatic variance/genotypic variance in the high LD population is lower than the ratio in the low LD population (Figure 2), regardless of the degree of inbreeding (30 to 60%). Note that both populations have the same average allele frequency (0.5). Compared to the populations with intermediate LD, the ratio epistatic variance/genotypic variance under high LD is greater relative to the non-inbred population with average allele frequency of 0.7 (approximately 10 to 60%) but lower relative to the other populations, regardless of the generation and degree of inbreeding (approximately 30 to 80%) (Figure 2).

Assuming an admixture of the types of epistasis and 30% of epistatic genes, the genetic variances in the non-inbred high LD population are higher than the values observed in the non-inbred low LD population (1.2 to 5.1 times higher). In the inbred populations, in general, the additive, additive x additive, and additive x dominance variances are greater under high LD but the dominance and the dominance x dominance variances are lower (Figures 1 and 2).

Discussion

WG Hill, ME Goddard and PM Visscher [10] emphasize that the knowledge about the relative magnitudes of the additive, dominance, and epistatic variances is important in evolutionary biology, medicine, and agriculture. However, the investigation about the joint significance of LD, epistasis, and inbreeding on the genetic variances for a quantitative trait is a challenge, even fixing a trait, i.e., even fixing the number of genes, the allele frequencies, and the degrees of dominance. One main

reason is that the theory available is too complex to allow the assessment of the relative magnitudes of the genetic variances [3, 4, 10, 19]. The other main reason is the large number of combinations between levels of LD (say, low to high) and inbreeding (say, not inbred to completely inbred) with distinct percentage of epistatic genes (say 30 to 100%), degree of epistasis (say, digenic to a high order), and type of epistasis (up to seven types of digenic epistasis, complementary or duplicate trigenic or high-order epistasis, or an admixture of types).

BS Weir and CC Cockerham [18] derived very complex functions for the components of the genotypic variance assuming a two-gene model with inbreeding, LD, and epistasis and concluded that they are of "little use". T Wang and ZB Zeng [19] only highlight that their theoretical results serve as a framework to understand and properly interpret estimates of the genetic effects and variance components in a QTL mapping experiment. The theoretical models investigated by WG Hill, ME Goddard and PM Visscher [10], assuming linkage equilibrium, predict high proportions of additive variance even in the presence of non-additive gene action. Assuming also linkage equilibrium, the theoretical results from A Maki-Tanila and WG Hill [4] showed that the epistatic variance is small compared to the additive variance, even assuming high heterozygosity. They also emphasize that the majority of the epistatic variance is due to two-locus interaction. Based on theoretical models including LD, WG Hill and A Maki-Tanila [3] confirmed that most of the genotypic variance in a segregating population is additive.

Because the main conclusion from the previously described studies is that most of the genotypic variance is additive, we believe that our simulation-based study provides significant additional knowledge about the influence of LD and epistasis on the genetic variances in non-inbred and inbred populations. Our study has a strong theoretical background on quantitative genetics. We assumed low to high LD levels for genes, not inbred to completely inbred populations, 30 and 100% of epistatic genes, and the seven types of digenic epistasis. Although there is evidence for high-order epistasis, pairwise interaction can contribute substantially to phenotypic variation between individuals [4, 20].

Our results agree with the main finds from WG Hill and A Maki-Tanila [3], A Maki-Tanila and WG Hill [4], and WG Hill, ME Goddard and PM Visscher [10], that LD significantly affects the genetic variances and that most of the genotypic variance is additive. However, from the analyses assuming an admixture of the types of epistasis and 30% of interacting genes, the ratio epistatic variance/genotypic variance was maximized in the populations with intermediate LD and average allele frequency of 0.3 (9 to 10%) and low LD and average allele frequency of 0.5 (10 to 22%), regardless of the generation and degree of inbreeding. The ratio was minimized in the populations with intermediate LD and average allele frequency of 0.7 (3 to 10%) and high LD and average allele frequency of 0.5 (3 to 8%). Our results also give support to the main conclusions of J Clo, J Ronfort and D Abu Awad [2], who assumed additive model under LD and distinct selfing rates. The differences observed for outcrossing species relies on their assumption of negative LD.

An important aspect to be also discussed is the unavailability of epistatic variance estimates from field phenotypic data. Most of the empirical evidence of epistasis comes from QTL mapping studies [9, 10] simply because when analyzing field data, there is no previous knowledge if there is epistasis. Further, even assuming digenic epistasis, linkage equilibrium, and non-inbred population, it would be necessary to estimate six independent variances and covariances between relatives to estimate the six genetic variances. Comparing estimates of the narrow and broad sense heritabilities only provides evidence of non-additive effects. Recently, however, some estimates of epistatic variances have been provided in studies involving genomic selection [21, 22]. In these studies, the epistatic variance ranged from 0 to 9.5% of the phenotypic variance.

Conclusions

Our main finds from a simulation-based study supported by quantitative genetics theory involving LD, epistasis, and inbreeding were: 1) the LD level for genes, even under a relatively low gene density, has a significant effect on the genetic variances in non-inbred and inbred populations; 2) assuming digenic epistasis, the additive variance is in general the most important component of the genotypic variance in non-inbred and inbred populations; 3) the ratio epistatic

variance/genotypic variance is proportional to the percentage of interacting genes and increases with random cross and selfing; 4) in general, the additive x additive variance is the most important component of the epistatic variance; and 5) the maximization of the epistatic variance depends on the allele frequency, level of LD, and epistasis type. Two important implications of our results are that selection based on breeding value prediction remains the best approach for population improvement and that cross- and self-pollinated populations keep a non-negligible amount of genetic variation for quantitative traits to allow their adaptive potential to environmental changes, assuming LD and epistasis.

Methods

Additive and dominance genetic values in inbred populations

Assume initially a single biallelic gene (A/a) determining a quantitative trait, where A is the gene that increases the trait expression, and a population derived by n generations of selfing from a Hardy-Weinberg equilibrium population (generation 0). Defining M_F^1 and M_F^2 as the means of the inbred population after an allelic substitution for the genes A and a, respectively, the average effect of the allelic genes in the inbred population are $\alpha_A^{(n)} = M_F^1 - M_F = q\alpha + 2Fpqd$ and $\alpha_a^{(n)} = M_F^2 - M_F = -p\alpha + 2Fpqd$, where $M_F = m + (p-q)a + 2pqd - 2Fpqd = M - 2Fpqd$ is the inbred population mean, p and q are the allelic frequencies, α is the average effect of an allelic substitution, F is the inbreeding coefficient, and M is the non-inbred population mean. Thus, the additive values in the inbred population are $A_{AA}^{(n)} = 2q\alpha + 4Fpqd = A_{AA}^{(0)} + 4Fpqd$, $A_{Aa}^{(n)} = (q-p)\alpha + 4Fpqd = A_{Aa}^{(0)} + 4Fpqd$, and $A_{aa}^{(n)} = -2p\alpha + 4Fpqd = A_{aa}^{(0)} + 4Fpqd$, where $A^{(0)}$ is the additive value in the non-inbred population. Note that $E\left(A^{(n)}\right) = 4Fpqd$. Expressing the genotypic values in the inbred population as a function of M_F, we have:

$$G_{AA} = M_F + A_{AA}^{(0)} + (-2q^2d + 2Fpqd) = M_F + A_{AA}^{(0)} + \left(D_{AA}^{(0)} + 2Fpqd\right) = M_F + A_{AA}^{(0)} + D_{AA}^{(n)}$$

$$G_{Aa} = M_F + A_{Aa}^{(0)} + (2pqd + 2Fpqd) = M_F + A_{Aa}^{(0)} + \left(D_{Aa}^{(0)} + 2Fpqd\right) = M_F + A_{Aa}^{(0)} + D_{Aa}^{(n)}$$

$$G_{aa} = M_F + A_{aa}^{(0)} + (-2p^2d + 2Fpqd) = M_F + A_{aa}^{(0)} + \left(D_{aa}^{(0)} + 2Fpqd\right) = M_F + A_{aa}^{(0)} + D_{aa}^{(n)}$$

Note that in the inbred population, $E\left(A^{(0)}\right) = E\left(D^{(n)}\right) = 0$ but $E\left(D^{(0)}\right) = -2Fpqd$. Note also that the additive value in the non-inbred population is the additive value in the inbred population expressed as deviation from its mean $\left(A^{(0)} = A^{(n)} - 4Fpqd\right)$ and the dominance value in the inbred population is the dominance value in the non-inbred population expressed as deviation from its mean $\left(D^{(n)} = D^{(0)} + 2Fpqd\right)$. This implies that, in the inbred population, $E(G) = M_F$.

Genetic variances in inbred populations in LD

Assume now two linked biallelic genes (A/a and B/b) determining a quantitative trait and a non-inbred population in LD (generation 0). Assume dominance but initially no epistasis. After n generations of selfing, the genotypic variance for the two genes in the inbred population is (see the genotype probabilities in the Appendix) $\sigma_G^{2(n)} = \sigma_A^{2(n)} + \sigma_D^{2(n)} + 2\sigma_{A,D}^{(n)}$, where:

$\sigma_A^{2(n)} = (1+F)(2p_aq_a\alpha_a^2 + 2p_bq_b\alpha_b^2) + 2[2 + c_1(1-2r_{ab})]\Delta_{ab}^{(-1)}\alpha_a\alpha_b = (1+F)\sigma_A^{2(0)} +$ $2[c_1(1-2r_{ab}) - 2F]\Delta_{ab}^{(-1)}\alpha_a\alpha_b$ is the additive variance,

$\sigma_D^{2(n)} = (1-F^2)(4p_a^2q_a^2d_a^2 + 4p_b^2q_b^2d_b^2) + F[4p_aq_a(p_a-q_a)^2d_a^2 + 4p_bq_b(p_b-q_b)^2d_b^2] +$ $8\left\{(1-F)(c^n - 1 + F)p_aq_ap_bq_b + (p_a-q_a)(p_b-q_b)[(1-F)c^n - (1-2F) + c_1(1-2r_{ab})/\right.$ $\left.2]\Delta_{ab}^{(-1)}/2 + (1-F)c^n{\Delta_{ab}^{(-1)}}^2\right\}d_ad_b = (1-F^2)\sigma_D^{2(0)} + FD_2 + 8\left\{(1-F)(c^n - 1 +\right.$ $\left.F)p_aq_ap_bq_b + (p_a-q_a)(p_b-q_b)[(1-F)c^n - (1-2F) + c_1(1-2r_{ab})/2]\Delta_{ab}^{(-1)}/2 +\right.$ $\left.[(1-F)c^n - (1-F^2)]{\Delta_{ab}^{(-1)}}^2\right\}d_ad_b$ is the dominance variance, and

$\sigma_{A,D}^{(n)} = 2F[2p_aq_a(p_a-q_a)\alpha_ad_a + 2p_bq_b(p_b-q_b)\alpha_bd_b] + [2F + c_1(1-2r_{ab})]\Delta_{ab}^{(-1)}[(p_b -$ $q_b)\alpha_ad_b + (p_a-q_a)\alpha_bd_a] = 2FD_1 + [2F + c_1(1-2r_{ab})]\Delta_{ab}^{(-1)}[(p_b - q_b)\alpha_ad_b +$ $(p_a-q_a)\alpha_bd_a]$ is the covariance between additive and dominance values,

where $\Delta_{ab}^{(-1)} = P_{AB}^{(-1)}.P_{ab}^{(-1)} - P_{Ab}^{(-1)}.P_{aB}^{(-1)}$ is the measure of LD in the gametic pool of generation –1 [23], where $P^{(-1)}$ is a haplotype probability, r_{ab} is the recombination frequency, $c_1 = 2\{1 - [(1-2r_{ab})/2]^n\}/(1+2r_{ab})$, $c = 1 - 2r_{ab}(1-r_{ab})$, $\sigma_A^{2(0)} = 2p_aq_a\alpha_a^2 + 2p_bq_b\alpha_b^2 +$

$4\Delta_{ab}^{(-1)}\alpha_a\alpha_b$ and $\sigma_D^{2(0)} = 4p_a^2q_a^2d_a^2 + 4p_b^2q_b^2d_b^2 + 8d_ad_b$ are the additive and dominance variances in the non-inbred population in LD [24], and D_1 (covariance of a and d) and D_2 (variance of d) are the components of the covariance of relatives from self-fertilization, assuming linkage equilibrium [8]. The other terms are the covariances between the average effects of an allelic substitution, between dominance deviations, and between the average effect of an allelic substitution and dominance deviation, for genes in LD. Because we assumed biallelic genes, $\breve{H} = \sigma_D^2$. Thus, $(1-F^2)\sigma_D^{2(0)} = (1-F)\sigma_D^{2(0)} + F(1-F)\breve{H}$. Note that the genotypic variance derived here is a general formulation for the Cockerham's genotypic variance c_{ggg} [8], assuming LD. If p = q, $\sigma_{A,D}^{(n)} = 0$.

Assuming LD but no inbreeding, the genotypic variance after n generations of random cross in the non-inbred population in LD is $\sigma_G^{2(n)} = \sigma_A^{2(n)} + \sigma_D^{2(n)}$, because $\sigma_{A,D}^{(n)} = 0$, where:

$$\sigma_A^{2(n)} = 2p_aq_a\alpha_a^2 + 2p_bq_b\alpha_b^2 + 4(1-r_{ab})^n\Delta_{ab}^{(-1)}\alpha_a\alpha_b$$

$$\sigma_D^{2(n)} = 4p_a^2q_a^2d_a^2 + 4p_b^2q_b^2d_b^2 + 8\left[(1-r_{ab})^n\Delta_{ab}^{(-1)}\right]^2 d_ad_b$$

Thus, the genotypic variance can increase or decreases after n generations of random cross in a non-inbred population, depending on the sign of the LD measure. The LD value is positive for genes in coupling phase and negative for genes in repulsion phase.

Epistasis in non-inbred and inbred populations in LD

The quantitative genetics theory for modelling epistasis in a population in LD is a generalization of the theory proposed by O Kempthorne [16], who assumed a non-inbred population in linkage equilibrium and any number of alleles. We assumed biallelism. It should be emphasized that the Kempthorne's theory allows a generalization from two to three or more interacting genes. But fitting three or more interacting genes in a population in LD is a challenge because the genotype probabilities for three or more genes in LD are too complex to derive. Furthermore, only complementary and duplicate epistasis can be easily defined for three or more epistatic genes.

Assume now that the two previous defined genes are epistatic. The genotypic value is [16]:

$G_{ijkl} = M + \alpha_i^1 + \alpha_j^1 + \alpha_k^2 + \alpha_l^2 + \delta_{ij}^1 + \delta_{kl}^2 + (\alpha^1\alpha^2)_{ik} + (\alpha^1\alpha^2)_{jk} + (\alpha^1\alpha^2)_{il} + (\alpha^1\alpha^2)_{jl} + (\alpha^1\delta^2)_{ikl} + (\alpha^1\delta^2)_{jkl} + (\delta^1\alpha^2)_{ijk} + (\delta^1\alpha^2)_{ijl} + (\delta^1\delta^2)_{ijkl} = M + A + D + AA + AD + DA + DD$

where AA, AD, DA, and DD are the additive x additive, additive x dominance, dominance x additive, and dominance x dominance epistatic genetic values.

The parametric values of the 36 parameters for the nine genotypic values are obtained by solving the equations $\beta = (X'VX)^{-1}X'Vy$, under the restrictions defined by O Kempthorne [16], where X is the incidence matrix, $V = diagonal\left\{f_{ij}^{(n)}\right\}$ is the diagonal matrix of the genotype probabilities, and y is the vector of the genotypic values $\left(G_{ij}\right)$ (i, j = 0, 1, and 2).

O Kempthorne [16] provided explicit functions for all effects because he assumed linkage equilibrium. Assuming LD makes very difficult to derive such functions but the following results hold:

1) the expectation of the breeding value is zero regardless of the degree of inbreeding in the population.

2) the expectation of the dominance value is $E(D)^{(n)} = p_a q_a F(\delta_{AA} - 2\delta_{Aa} + \delta_{aa}) + p_b q_b F(\delta_{BB} - 2\delta_{Bb} + \delta_{bb})$; then, defining the dominance value in an inbred population as the dominance value expressed as deviation from its mean $\left(D^{(n)} = D - E(D)^{(n)}\right)$, $E\left(D^{(n)}\right) = 0$.

3) the expectation of the additive x additive value is zero only if there is no LD.

4) the expectation of the additive x dominance value is zero only if F = 0 or p = q for all genes.

5) the expectation of the dominance x additive value is zero only if F = 0 or p = q for all genes.

6) the expectation of the dominance x dominance value is zero only if F = 0 and there is no LD.

Thus, defining the additive x additive, additive x dominance, dominance x additive, and dominance x dominance epistatic values as the values expressed as deviation from its mean,

$AA^{(n)} = AA - E(AA)^{(n)}$, $AD^{(n)} = AD - E(AD)^{(n)}$, $DA^{(n)} = DA - E(DA)^{(n)}$, and $DD^{(n)} = DD - E(DD)^{(n)}$, the genotypic value in an inbred population can be expressed as

$$G = M + E(D)^{(n)} + E(AA)^{(n)} + E(AD)^{(n)} + E(DA)^{(n)} + E(DD)^{(n)} + A + D^{(n)} + AA^{(n)} + AD^{(n)} + DA^{(n)} + DD^{(n)} = M_F + A + D^{(n)} + AA^{(n)} + AD^{(n)} + DA^{(n)} + DD^{(n)}$$

This implies that $E(G) = M_F$. If F = 0 then

$$G = M + E(AA) + E(DD) + A + D + [AA - E(AA)] + AD + DA + [DD - E(DD)] = M^* + A + D + AA^* + AD + DA + DD^*$$

where,

$$E(AA) = 2\Delta_{ab}^{(-1)}(\alpha_A\alpha_B - \alpha_A\alpha_b - \alpha_a\alpha_B + \alpha_a\alpha_b) \quad \text{and} \quad E(DD) = \left[\Delta_{ab}^{(-1)}\right]^2(\delta_{AA}\delta_{BB} - 2\delta_{AA}\delta_{Bb} + \delta_{AA}\delta_{bb} - 2\delta_{Aa}\delta_{BB} + 4\delta_{Aa}\delta_{Bb} - \delta_{Aa}\delta_{bb} + \delta_{aa}\delta_{BB} - 2\delta_{aa}\delta_{Bb} + \delta_{aa}\delta_{bb}).$$

This implies that $E(G) = M^*$. If F = 0 and there is no LD,

$$G = M + A + D + AA + AD + DA + DD$$

where the linear components are those defined by O Kempthorne [16]. This implies that $E(G) = M$.

In non-inbred populations in LD, only the additive and dominance values are not correlated. The genotypic variance in these populations is, in simplified form,

$$\sigma_G^{2(0)} = \sigma_A^{2(0)} + \sigma_D^{2(0)} + \sigma_{AA}^{2(0)} + 2\sigma_{A,AA}^{(0)} + 2\sigma_{D,AA}^{(0)} + \cdots$$

where

$$\sigma_{AA}^{2(0)} = f_{22}^{(0)}[(4\alpha_A\alpha_B)]^2 + \cdots + f_{00}^{(0)}[(4\alpha_a\alpha_b)]^2 - \left[E(AA)^{(0)}\right]^2$$

$$\sigma_{A,AA}^{(0)} = 2\Delta_{ab}^{(-1)}[\alpha^A(\alpha_A\alpha_B - \alpha_A\alpha_b + \alpha_a\alpha_B - \alpha_a\alpha_b) + \alpha^B(\alpha_A\alpha_B - \alpha_a\alpha_B + \alpha_A\alpha_b - \alpha_a\alpha_b)]$$

$$\sigma_{D,AA}^{(0)} = -4\Delta_{ab}^{(-1)}[p_aq_ad_a(\alpha_A\alpha_B - \alpha_A\alpha_b - \alpha_a\alpha_B + \alpha_a\alpha_b) + p_bq_bd_b(\alpha_A\alpha_B - \alpha_a\alpha_B - \alpha_A\alpha_b + \alpha_a\alpha_b)]$$

where, to avoid confusion, α^A and α^B are the average effects of an allelic substitution.

The assumption of LD makes very difficult to derive the components of the genotypic variance (additive, dominance, and epistatic variances and the covariances between these effects),

even assuming non-inbred populations, biallelic genes, and only digenic epistasis. In respect to the types of digenic epistasis, the following can be defined [25, 26]:

1. Complementary ($G_{22} = G_{21} = G_{12} = G_{11}$ and $G_{20} = G_{10} = G_{02} = G_{01} = G_{00}$; proportion of 9:7 in a F_2).

2. Duplicate ($G_{22} = G_{21} = G_{20} = G_{12} = G_{11} = G_{10} = G_{02} = G_{01}$; proportion of 15:1 in a F_2).

3. Dominant ($G_{22} = G_{21} = G_{20} = G_{12} = G_{11} = G_{10}$ and $G_{02} = G_{01}$; proportion of 12:3:1 in a F_2).

4. Recessive ($G_{22} = G_{21} = G_{12} = G_{11}$, $G_{02} = G_{01}$, and $G_{20} = G_{10} = G_{00}$; proportion of 9:3:4 in a F_2)

5. Dominant and recessive ($G_{22} = G_{21} = G_{12} = G_{11} = G_{20} = G_{10} = G_{00}$ and $G_{02} = G_{01}$; proportion of 13:3 in a F_2).

6. Duplicate genes with cumulative effects ($G_{22} = G_{21} = G_{12} = G_{11}$, and $G_{20} = G_{10} = G_{02} = G_{01}$; proportion of 9:6:1 in a F_2).

7. Non-epistatic genic interaction ($G_{22} = G_{21} = G_{12} = G_{11}$, $G_{20} = G_{10}$, and $G_{02} = G_{01}$; proportion of 9:3:3:1 in a F_2).

Simulated data sets

Because the magnitude of the components of the genotypic variance generally cannot be inferred from the previous functions, all means and genetic variances and covariances were computed from simulated data sets provided by the software *REALbreeding* (available upon request). This software uses the quantitative genetics theory that was described in the previous sections and in JMS Viana [24]. *REALbreeding* has been used to provide simulated data in investigations in the areas of genomic selection [27], GWAS [28], QTL mapping [29], linkage disequilibrium [30], population structure [31], and heterotic grouping/genetic diversity [32].

The software simulates individual genotypes for genes and molecular markers and phenotypes in three steps using user inputs. The first step (genome simulation) is the specification of the number of chromosomes, molecular markers, and genes as well as marker type and density. The second step (population simulation) is the specification of the population(s) and sample size or

progeny number and size. A population is characterized by the average frequency for the genes (biallelic) and markers (first allele). The final step (trait simulation) is the specification of the individual phenotypes. In this stage, the user informs the minimum and maximum genotypic values for homozygotes (to compute the a deviations), the minimum and maximum phenotypic values (to avoid outliers), the direction and degree of dominance (to compute the dominance deviations/d), and the broad sense heritability. The current version allows the inclusion of digenic epistasis, gene x environment interaction, and multiple traits (up to 10), including pleiotropy. The population mean (M), additive (A), dominance (D), and epistatic (AA, AD, DA, and DD) genetic values or general and specific combining ability effects (GCA and SCA) or genotypic values (G) and epistatic values (I), depending on the population, are calculated from the parametric gene effects and frequencies and the parametric LD values. The phenotypic values (P) are computed assuming error effects (E) sampled from a normal distribution ($P = M + A + D + AA + AD + DA + DD + E = G + E$ or $P = M + GCA1 + GCA2 + SCA + I + E = G + E$). The population in LD is generated by crossing two populations in linkage equilibrium followed by a generation of random cross. This generation of random cross aims to generate a population in Hardy-Weinberg equilibrium. Thus, the generation 0 (the founder population) is a population in Hardy-Weinberg equilibrium, in LD for linked genes and molecular markers, and the individuals are not related. The parametric LD in this population is $\Delta_{ab}^{(-1)} = [(1 - 2r_{ab})/4](p_{a1} - p_{a2})(p_{b1} - p_{b2})$, where the indexes 1 and 2 stand for the allele frequencies in the parental populations.

The quantitative genetics theory for epistasis does not solve the challenge of studying genetic variability and covariance between relatives in populations, using simulated data sets, even assuming simplified scenarios such as linkage equilibrium and no inbreeding. Because the genotypic values for any two interacting genes are not known, there are infinite genotypic values that satisfy the specifications of each type of digenic epistasis. For example, fixing the gene frequencies (the population) and the parameters m, a, d, and d/a (degree of dominance) for each gene (the trait), the solutions $G_{22} = G_{21} = G_{12} = G_{11} = 5.25$ and $G_{20} = G_{10} = G_{02} = G_{01} = G_{00} =$

5.71 or $G_{22} = G_{21} = G_{12} = G_{11} = 6.75$ and $G_{20} = G_{10} = G_{02} = G_{01} = G_{00} = 2.71$ define complementary epistasis but the genotypic values are not the same.

The solution implemented in the software allows the user to control the magnitude of the epistatic variance (V(I)), relative to the magnitudes of the additive and dominance variances (V(A) and V(D)). As an input for the user, the software requires the ratio V(I)/(V(A) + V(D)) for each pair of interacting genes (a single value; for example, 1.0). Then, for each pair of epistatic genes the software samples a random value for the epistatic value I_{22} (the epistatic value for the genotype AABB), assuming $I_{22} \sim N(0, V(I))$. Then, the other epistatic effects and genotypic values are computed.

We simulated grain yield assuming 400 genes in 10 chromosomes of 200 and 50 cM (40 genes/chromosome). The average density was approximately one gene each five and one cM, respectively. We generated five populations, two with high LD level and one with low LD level, all three with an average allele frequency of 0.5, and two populations with intermediate LD level and an average frequency for the favorable genes of 0.3 (not improved) and 0.7 (improved). We defined positive dominance (average degree of dominance of 0.6), maximum and minimum genotypic values for homozygotes of 160 and 30 $g.plt^{-1}$, and maximum and minimum phenotypic values of 180 and 10 $g.plt^{-1}$. The broad sense heritability was 20%. For each population we assumed additive-dominance model and additive-dominance with digenic epistasis model, defining 100% and 30% of interacting genes. Concerning the ratio V(I)/(V(A) + V(D)), the analyses assuming ratios 1, 10, and 100 evidenced that increasing the ratio from 1 to 10 and 100 increased the epistatic variances but also increased the additive and dominance variances. Then, because the main conclusions for the greater ratios were essentially the same provided by ratio 1, we will present only the results for ratio 1. With epistasis, we assumed a single type or an admixture of the seven types. We ranged the degree of inbreeding from 0.0 to 1.0, assuming 10 generations of selfing. We also assumed 10 generations of random crosses. The population size was 5,000 per generation.

The characterization of the LD in the populations was based on the parametric Δ, r^2, and D' values for the 40 genes in chromosome 1, which were provided by *REALbreeding* (it should be similar for the other chromosomes). The heatmaps were processed using the R package pheatmap. Assuming no epistasis, the software provides the parametric additive and dominance genetic values and the parametric genetic variances and covariances. Assuming epistasis, the software provides the parametric additive, dominance, and epistatic genetic values. Thus, under epistasis, the genetic variances and covariances were computed from the parametric genetic values, using a sample size of 5,000 individuals per generation. Two important implications of our results are that selection based on breeding value prediction remains the best approach for population improvement and that cross- and self-pollinated populations keep a non-negligible amount of genetic variation for quantitative traits to allow their adaptive potential to environmental changes, assuming LD and epistasis.

List of Abbreviations: LD – linkage disequilibrium; A – additive value; D – dominance value; AA – additive x additive value; AD – additive x dominance value; AD – dominance x additive value; DD – dominance x dominance value; G – genotypic value; I – epistatic value.

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: The data set is available at https://doi.org/10.6084/m9.figshare.13607306.

Competing Interests: The authors declare that they have no competing interests.

Funding: None.

Authors' contributions: JMSV designed the study, developed the software, processed the data, and wrote the manuscript. AAFG designed the study, processed the data, and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements: We thank the National Council for Scientific and Technological Development (CNPq), the Brazilian Federal Agency for Support and Evaluation of Graduate Education (Capes; Finance Code 001), and the Foundation for Research Support of Minas Gerais State (Fapemig) for financial support.

References

1. Fisher RA: **The correlation between relatives on the supposition of Mendelian inheritance**. *Transactions of the Royal Society of Edinburgh* 1918, **52**(2):399-433.
2. Clo J, Ronfort J, Abu Awad D: **Hidden genetic variance contributes to increase the short-term adaptive potential of selfing populations**. *Journal of Evolutionary Biology* 2020, **33**(9):1203-1215.
3. Hill WG, Maki-Tanila A: **Expected influence of linkage disequilibrium on genetic variance caused by dominance and epistasis on quantitative traits**. *Journal of Animal Breeding and Genetics* 2015, **132**(2):176-186.
4. Maki-Tanila A, Hill WG: **Influence of Gene Interaction on Complex Trait Variation with Multilocus Models**. *Genetics* 2014, **198**(1):355-367.
5. Clo J, Gay L, Ronfort J: **How does selfing affect the genetic variance of quantitative traits? An updated meta-analysis on empirical results in angiosperm species**. *Evolution* 2019, **73**(8):1578-1590.
6. Hasselgren M, Noren K: **Inbreeding in natural mammal populations: historical perspectives and future challenges**. *Mammal Review* 2019, **49**(4):369-383.
7. Howard JT, Pryce JE, Baes C, Maltecca C: **Invited review: Inbreeding in the genomics era: Inbreeding, inbreeding depression, and management of genomic variability**. *Journal of dairy science* 2017, **100**(8):6009-6024.
8. Cockerham CC: **Covariances of relatives from self-fertilization** *Crop Science* 1983, **23**:1177-1180.

9. Mackay TFC: **Epistasis and quantitative traits: using model organisms to study gene-gene interactions**. *Nature Reviews Genetics* 2014, **15**(1):22-33.

10. Hill WG, Goddard ME, Visscher PM: **Data and theory point to mainly additive genetic variance for complex traits**. *Plos Genetics* 2008, **4**(2).

11. Vitezica ZG, Legarra A, Toro MA, Varona L: **Orthogonal Estimates of Variances for Additive, Dominance, and Epistatic Effects in Populations**. *Genetics* 2017, **206**(3):1297-1307.

12. Forneris NS, Vitezica ZG, Legarra A, Perez-Enciso M: **Influence of epistasis on response to genomic selection using complete sequence data**. *Genetics Selection Evolution* 2017, **49**.

13. Su G, Christensen OF, Ostersen T, Henryon M, Lund MS: **Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers**. *PloS one* 2012, **7**(9):e45293.

14. Monir MM, Zhu J: **Comparing GWAS Results of Complex Traits Using Full Genetic Model and Additive Models for Revealing Genetic Architecture**. *Scientific Reports* 2017, **7**.

15. Misztal I, Aguilar I, Lourenco D, Ma L, Steibel J, Toro M: **Emerging issues in genomic selection**. *Journal of animal science* 2021.

16. Kempthorne O: **The theoretical values of correlations between relatives in random mating populations**. *Genetics* 1954, **40**:153-167.

17. Cockerham CC: **An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present**. *Genetics* 1954(39):859-882.

18. Weir BS, Cockerham CC: **Two-locus theory in quantitative genetics**. In: *International Conference on Quantitative Genetics: 1976; Ames*. The Iowa State University Press; 1977: 247-269.

19. Wang T, Zeng ZB: **Models and partition of variance for quantitative trait loci with epistasis and linkage disequilibrium**. *BMC genetics* 2006, **7**.

20. Domingo J, Baeza-Centurion P, Lehner B: **The Causes and Consequences of Genetic Interactions (Epistasis)**. *Annual Review of Genomics and Human Genetics, Vol 20, 2019* 2019, **20**:433-460.

21. Chen ZQ, Baison J, Pan J, Westin J, Gil MRG, Wu HX: **Increased Prediction Ability in Norway Spruce Trials Using a Marker X Environment Interaction and Non-Additive Genomic Selection Model**. *Journal of Heredity* 2019, **110**(7):830-843.

22. Vitezica ZG, Reverter A, Herring W, Legarra A: **Dominance and epistatic genetic variances for litter size in pigs using genomic models**. *Genetics Selection Evolution* 2018, **50**.

23. Kempthorne O: **An Introduction to Genetic Statistics**. Ames: The Iowa State University Press; 1973.

24. Viana JMS: **Quantitative genetics theory for non-inbred populations in linkage disequilibrium**. *Genetics and Molecular Biology* 2004, **27**(4):594-601.

25. Viana JMS: **Dominance, epistasis, heritabilities and expected genetic gains**. *Genetics and Molecular Biology* 2005, **28**(1):67-74.

26. Viana JMS: **Components of variation of polygenic systems with digenic epistasis**. *Genetics and Molecular Biology* 2000, **23**(4):883-892.

27. Viana JMS, Pereira HD, Piepho HP, Silva FFE: **Efficiency of Genomic Prediction of Nonassessed Testcrosses**. *Crop Science* 2019, **59**(5):2020-2027.

28. Pereira HD, Viana JMS, Andrade ACB, Silva FFE, Paes GP: **Relevance of genetic relationship in GWAS and genomic prediction**. *Journal of Applied Genetics* 2018, **59**(1):1-8.

29. Viana JMS, Silva FF, Mundim GB, Azevedo CF, Jan HU: **Efficiency of low heritability QTL mapping under high SNP density**. *Euphytica* 2017, **213**(1).

30. Andrade ACB, Viana JMS, Pereira HD, Pinto VB, Fonseca ESF: **Linkage disequilibrium and haplotype block patterns in popcorn populations**. *PloS one* 2019, **14**(9):e0219417.

31. Viana JMS, Valente MSF, Silva FF, Mundim GB, Paes GP: **Efficacy of population structure analysis with breeding populations and inbred lines**. *Genetica* 2013, **141**(7-9):389-399.

32. Viana JMS, Risso LA, Oliveira deLima R, Fonseca e Silva F: **Factors affecting heterotic grouping with cross-pollinating crops**. *Agronomy Journal* 2020.

Appendix

For two genes, the genotype probabilities in generation 0 ($f_{ij}^{(0)}$) are presented by JMS Viana [24], where i and j (i, j = 0, 1, or 2) are the number of copies of the gene that increase the trait expression (A and B). For example, $f_{22}^{(0)} = p_a^2 p_b^2 + 2p_a p_b \Delta_{ab}^{(-1)} + \left[\Delta_{ab}^{(-1)}\right]^2$. After n generations of selfing, the genotype probabilities are:

$$f_{22}^{(n)} = f_{22}^{(0)} + (F/2)\left[f_{21}^{(0)} + f_{12}^{(0)}\right] + P_1^{(n)}$$

$$f_{21}^{(n)} = (1-F)\left[f_{21}^{(0)} + (1-c^n)f_{11}^{(0)}/2\right]$$

$$f_{20}^{(n)} = f_{20}^{(0)} + (F/2)\left[f_{21}^{(0)} + f_{10}^{(0)}\right] + P_2^{(n)}$$

$$f_{12}^{(n)} = (1-F)\left[f_{12}^{(0)} + (1-c^n)f_{11}^{(0)}/2\right]$$

$$f_{11}^{(n)} = (1-F)c^n f_{11}^{(0)}$$

$$f_{10}^{(n)} = (1-F)\left[f_{10}^{(0)} + (1-c^n)f_{11}^{(0)}/2\right]$$

$$f_{02}^{(n)} = f_{02}^{(0)} + (F/2)\left[f_{01}^{(0)} + f_{12}^{(0)}\right] + P_2^{(n)}$$

$$f_{01}^{(n)} = (1-F)\left[f_{01}^{(0)} + (1-c^n)f_{11}^{(0)}/2\right]$$

$$f_{00}^{(n)} = f_{00}^{(0)} + (F/2)\left[f_{01}^{(0)} + f_{10}^{(0)}\right] + P_1^{(n)}$$

where

$$P_1^{(n)} = (1/4)\left\{[F-(1-F)(1-c^n)]f_{11}^{(0)} + c_1(1-2r_{ab})\Delta_{ab}^{(-1)}\right\}$$

$$P_2^{(n)} = (1/4)\left\{[F-(1-F)(1-c^n)]f_{11}^{(0)} - c_1(1-2r_{ab})\Delta_{ab}^{(-1)}\right\}$$

$$c = 1 - 2r_{ab}(1-r_{ab})$$

$$c_1 = 2\{1-[(1-2r_{ab})/2]^n\}/(1+2r_{ab})$$

and r_{ab} is the recombination frequency.

File name: Additional File
Format: .pdf
Title of data: Figures
Description of data: Additional figures

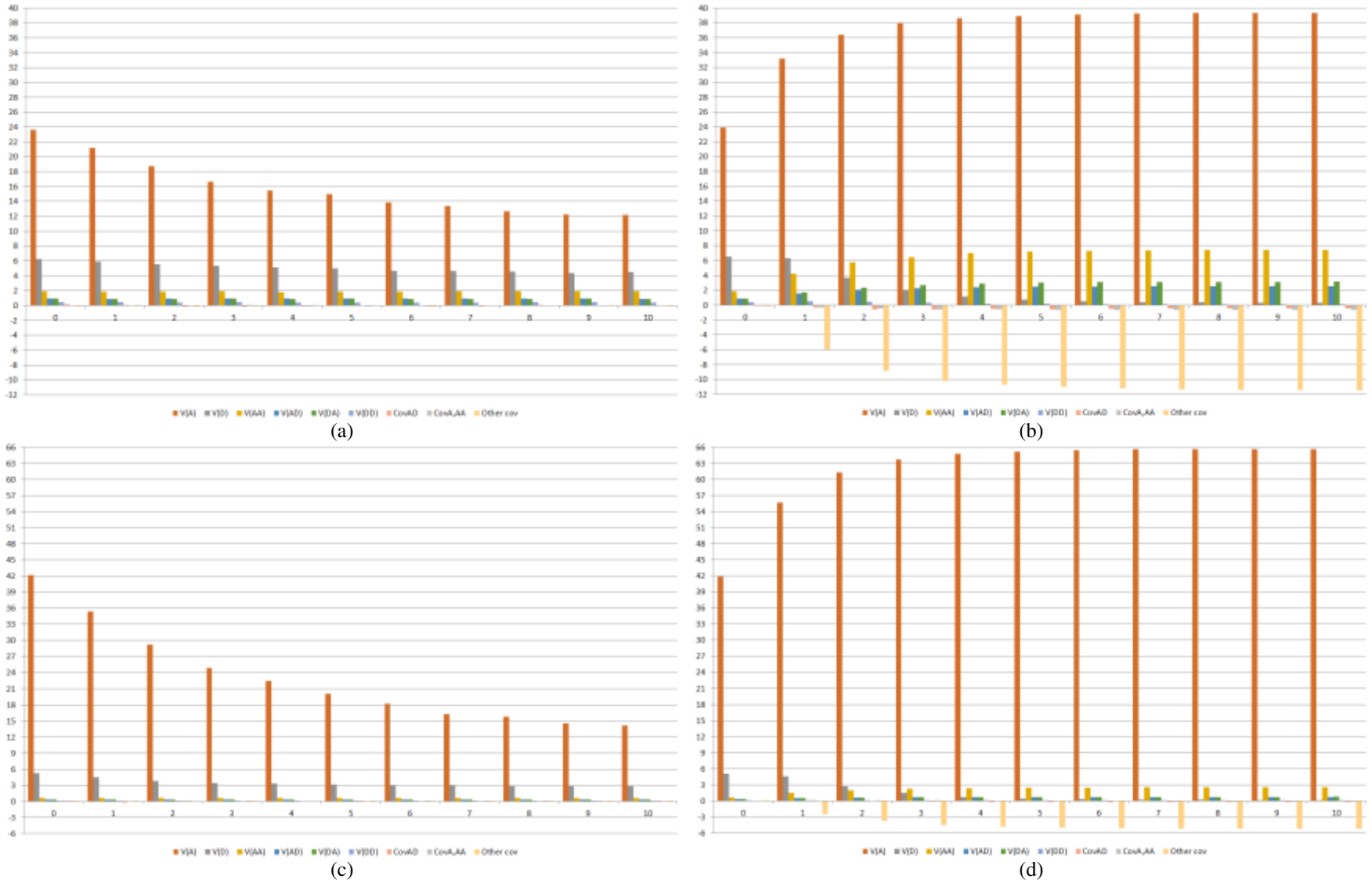
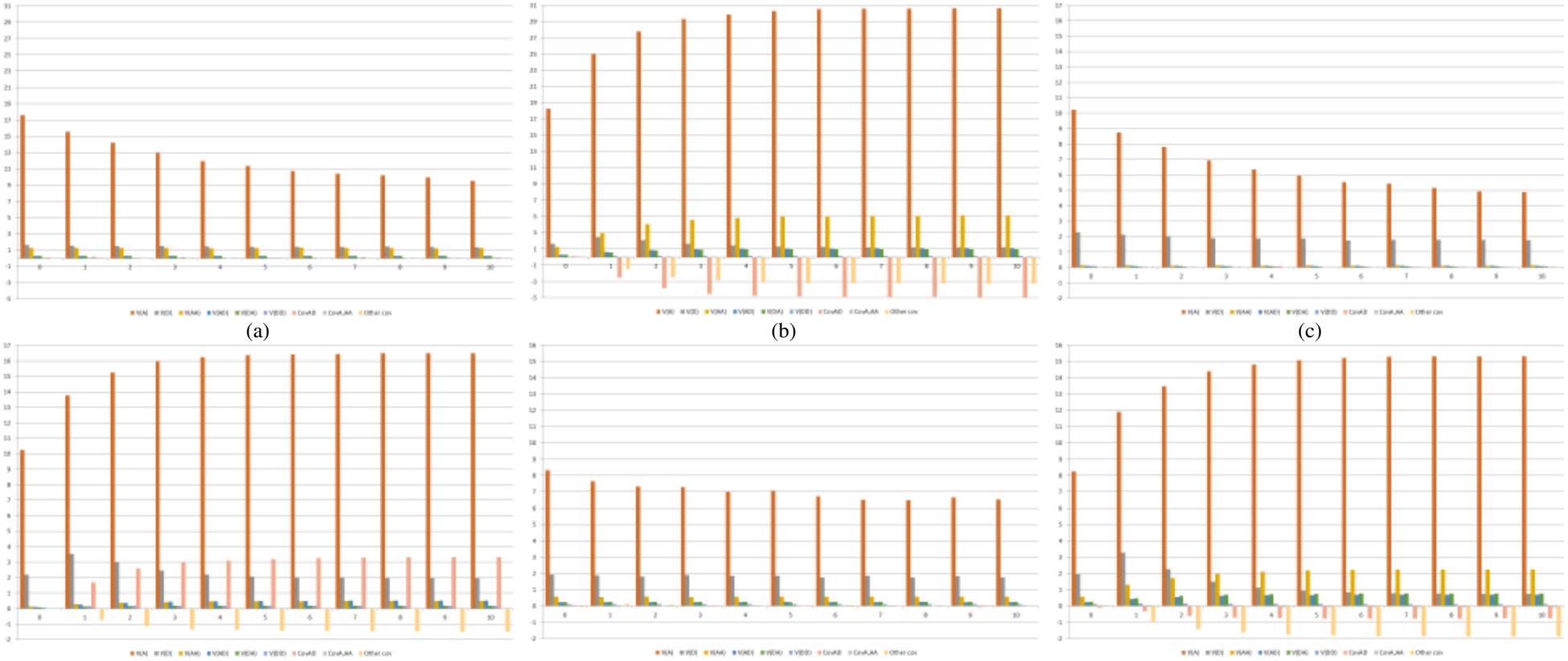


Figure 1. Components of the genotypic variance in population with high LD level, along 10 generations of random cross (a and c) or selfing (b and d), assuming an admixture of digenic epistasis, 100 (a and b) and 30% (c and d) of epistatic genes, and sample size of 5,000 per generation.



(a) (b) (c)

(d) (e) (f)

Figure 2. Components of the genotypic variance in the populations not improved (a and b) and improved (c and d), with intermediate LD level, and in the population with low LD level (e and f), along 10 generations of random cross (a, c, and e) or selfing (b, d, and f), assuming an admixture of digenic epistasis, 30% of epistatic genes, and sample size of 5,000 per generation.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile.pdf