Received: 26 September 2018 | Revised: 23 December 2018 | Accepted: 23 January 2019

DOI: 10.1111/jeb.13421

RESEARCH PAPER

WILEY Journal of Evolutionary Biology

Inferring the potentially complex genetic architectures of adaptation, sexual dimorphism and genotype by environment interactions by partitioning of mean phenotypes

Andrew Armstrong[1] | Nathan W. Anderson[1] | Heath Blackmon[2]

[1]Department of Mathematics, Texas A&M University, College Station, Texas

[2]Department of Biology, Texas A&M University, College Station, Texas

Correspondence
Heath Blackmon, Department of Biology, Texas A&M University, College Station, TX.
Email: coleoguy@gmail.com

Abstract

Genetic architecture fundamentally affects the way that traits evolve. However, the mapping of genotype to phenotype includes complex interactions with the environment or even the sex of an organism that can modulate the expressed phenotype. Line-cross analysis is a powerful quantitative genetics method to infer genetic architecture by analysing the mean phenotype value of two diverged strains and a series of subsequent crosses and backcrosses. However, it has been difficult to account for complex interactions with the environment or sex within this framework. We have developed extensions to line-cross analysis that allow for gene by environment and gene by sex interactions. Using extensive simulation studies and reanalysis of empirical data, we show that our approach can account for both unintended environmental variation when crosses cannot be reared in a common garden and can be used to test for the presence of gene by environment or gene by sex interactions. In analyses that fail to account for environmental variation between crosses, we find that line-cross analysis has low power and high false-positive rates. However, we illustrate that accounting for environmental variation allows for the inference of adaptive divergence, and that accounting for sex differences in phenotypes allows practitioners to infer the genetic architecture of sexual dimorphism.

KEYWORDS

environmental effect, epistasis, G × E, line-cross analysis, sexual dimorphism

1 | INTRODUCTION

Identifying the genetic architecture underlying traits is a central goal of evolutionary biologists (Demuth, Flanagan, & Delph, 2014; Fuchsberger et al., 2016; Kitano et al., 2009; Küpper et al., 2016), as well as animal and plant breeders (Gall, 1975; Pooni, Jinks, & de Toledo, 1985; Singh, Bhullar, & Gill, 1986). The genetic architecture of a trait may, at its simplest, be described by the action of a single locus whose effect can be decomposed into additive and dominance components. However, rarely does this simple view describe traits of interest to biologists. For instance, empirical studies have found that life-history traits, complex diseases and hybrid incompatibilities are all frequently impacted by complex genetic architectures that include epistasis (i.e. interactions among loci) (Demuth & Wade, 2007; Pandey et al., 2012; Roff & Emerson, 2006). The importance of these epistatic interactions where the impact of an allele at one locus is dependent on the genetic background (i.e. alleles present at other loci) has been a topic of contention since the dawn of the modern synthesis (Fisher, 1958; Wright, 1931). Many of the possible roles of epistatic

Data deposited at Dryad: https://doi:10.5061/dryad.fd0s01b

variation in the evolutionary dynamics of traits are clear mathematically (Goodnight, 1988; Wade & Goodnight, 1998; Wolf, Brodie, Cheverud, Moore, & Wade, 1998). However, substantial disagreement remains about how important epistasis is in nature and the course of adaptation and speciation (Cheverud & Routman, 1996; Coyne, Barton, & Turelli, 1997; Peck, Ellner, & Gould, 1998; Turelli & Barton, 2006). One cause of this unresolved debate is variation in the methods we use to infer genetic architecture.

One of the first and most widely used methods of exploring genetic architecture is the partitioning of additive and dominance variance based on phenotypic variance and covariance of relatives (Falconer & Mackay, 1989). This approach has been successful in predicting short-term response to selection within lines or strains, but variance partitioning of this type is not well designed to estimate epistatic effects (Falconer & Mackay, 1989). With the advent of QTL and later GWAS, explicit estimates of epistatic effects became possible. However, while technological and computational advancements have allowed the application of GWAS and QTL studies to a continually increasing number of systems, these approaches are not without their own set of problems (Visscher et al., 2017). For instance, even if we are only interested in additive effects, these methods are tasked with performing thousands of tests with only a limited sample size. This statistical challenge introduces problems with regard to false positives, bias in types of architecture investigated and the effect size necessary for reliable inference (Donnelly, 2008; Korte & Farlow, 2013; Rockman, 2012; Wei, Hemani, & Haley, 2014).

QTL and GWAS methods were developed to identify the location of genes that impact a trait. In contrast, line-cross analysis (LCA) is an alternative method that dispenses with the goal of locus identification and instead focuses on identification of the trait architecture or mode of gene action (i.e. additive, dominance, epistasis) that underlies a trait. In LCA, two parental strains are crossed creating an F1. This F1 is then used to generate subsequent backcrosses. Depending on the genetic architecture of interest, reciprocal crosses may also be generated. In each of these crosses, or cohorts, the phenotype is measured and we are able to partition differences in the means of cohorts into additive, dominance and epistatic components (Cavalli, 1952; Hayman, 1958; Mather & Jinks, 1982). With this approach, we dispense with the goal of identifying nucleotides or genes of interest and instead focus on the net composite genetic effect (CGE) of all loci that influence a trait. In many LCA experiments, the focus is on simple CGEs (e.g. additive, dominance and the three types of interactions among these). These interactions: additive by additive, additive by dominance and dominance by dominance can capture variation in line means best explained by different patterns of epistatic interactions (Brodie, Wade, & Wolf, 2000). However, with adequately designed experiments, much more complex genetic architectures can be investigated. For instance, if males and females are measured separately or in equal numbers, additive and dominance effects may be further partitioned into either autosomal additive, autosomal dominance, X chromosome additive, X chromosome dominance and Y chromosome additive (Demuth et al., 2014; Miller, Starmer, & Pitnick, 2003). Although we discuss XY systems, LCA and our implementation thereof is capable of dealing with XY or ZW sex chromosome systems as well as systems without sex chromosomes (Blackmon & Demuth, 2016). One weakness of LCA is that it estimates a net effect, meaning that loci with the same mode of gene action (e.g. additive) but with opposite direction of effect will cancel out and reduce the signal for that CGE. This should make LCA inference conservative but prone to type II errors. In the past, LCA was hampered due to the sheer number of biologically realistic potential combinations of genetic effects and the computational and model selection problems this created (Demuth & Wade, 2006; Mundry & Nunn, 2008). However, a recently developed information theoretic approach with model averaging can make LCA a powerful and statistically conservative tool for inferring complex genetic architectures (Blackmon & Demuth, 2016; van Heerwaarden & Sgrò, 2017).

Despite being developed over six decades ago and being extended recently, several challenges continue to limit the application of LCA to empirical data. (a) The construction of a C-matrix (a matrix of coefficients describing the opportunity for CGEs to impact the mean phenotype of a cohort) is the cornerstone of an LCA (Demuth & Wade, 2006). However, the construction of the C-matrix can be a challenge due to its sheer size and can bias downstream results if only some of the possible CGEs are included. (b) While it has long been recognized that environmental interactions could be incorporated into an LCA framework, no existing software provides practitioners with a simple way to do this (Bulmer, 1980; Rundle & Whitlock, 2001), and this is despite broad interest in G × E interactions (Egan & Funk, 2009; Rundle, 2002). (c) No existing framework is available to incorporate both male and female data in analyses that focus on traits that are sexually dimorphic. Despite abundant empirical evidence that the sex of an individual and the alleles, it carries can interact and potentially result in sexual dimorphism (Ledón-Rettig, Zattara, & Moczek, 2017; Ober, Loisel, & Gilad, 2008; Poissant, Wilson, & Coltman, 2010; Weiss, Pan, Abney, & Ober, 2006; Wolak, 2013). (d) Finally, we lack a broad understanding of the potential biases that may result from LCA in the presence of unintended environmental variation (e.g. differences in temperature or humidity within an incubator or greenhouse).

We solved these problems by extending traditional LCA approaches in several ways (implemented in the open source R package SAGA), conducting extensive simulation studies and reanalysing empirical data. We solve the difficulty of C-matrix construction by providing users with a simple data input format (Table 1) where users describe their cohorts and the software applies an algorithm to construct the C-matrix appropriate for the experiment. By extending the construction of the C-matrix to include an environmental effect as first described by Bulmer (1980), we have removed the previous barrier to analyses of gene by environment interactions. Likewise, we use a similar approach to account for sex by gene interactions facilitating the analysis of sexual dimorphism. Finally, we use this improved version of the R package SAGA to perform extensive simulation testing and reanalysis of published data to evaluate the perils and promise of the inclusion of environmental and sex effects. Briefly, we show that

the inclusion of an environmental variable can greatly increase the power and versatility of LCA to uncover the genetic architecture of traits without inflating false-positive rates. Furthermore, we show the potential danger of undocumented environmental variation. We find that among-cohort environmental variation can lead to excessive false positives. Finally, we show that the same approach that we use to incorporate environmental variation can also be successfully applied to reveal the genetic architecture of adaptive divergence and sexual dimorphism. Taken together, the theoretical extensions that we provide in our R package SAGA should make LCA a more flexible and easily used tool for evolutionary biologists. Our simulation results underscore the great importance of careful experimental design to eliminate among-family differences in environmental conditions. When this is impossible, the extensions we provide allow researchers to include a measured or even unmeasured environmental variable that may be confounding the analysis of genetic architecture.

2 | MATERIALS AND METHODS

2.1 | Weighted least squares

The math that underlies an LCA is a weighted least squares regression, and we can represent our model of the genetic architecture as the linear model (1).

$$y = C\beta + e \tag{1}$$

Here, y represents the vector of observed cohort means, C is the matrix that describes the opportunity for each CGE to impact the phenotype of a cohort, β is the vector of parameters to be estimated that describe the degree to which each CGE is responsible for the observed cohort means, and e is a vector of the random errors associated with the means of each cohort. In the weighted least squares approach, we then find the estimate of the parameters $\hat{\beta}$ that minimizes the weighted sum of squares (2).

$$(y - C\beta)^T V^{-1} (y - C\beta) \tag{2}$$

Here, V is the variance–covariance matrix of e. In LCA, V is a diagonal matrix with the standard errors of cohort means along the diagonal. This scales each cohort's contribution to the sum of squares by the certainty of the cohort mean. When few or large effect loci are responsible for differences in the phenotype of interest, variation in relatedness among members of a cohort could increase variance and reduce certainty in the cohort mean. The less certain a cohort mean, the less the contribution from that cohort to the sum of squares. The parameter estimates $\hat{\beta}$ are

$$\hat{\beta} = (C^T V^{-1} C)^{-1} C^T V^{-1} y \tag{3}$$

In our implementation, we use the R function `glm` (part of the core R stats package) to perform the maximum-likelihood weighted least squares regression and calculate the AIC of each possible model. A detailed description of the process of using these results in multimodel inference has been discussed in depth in the past (Blackmon & Demuth, 2016).

TABLE 1 Example of a typical data set that a user would supply. The cohort column describes the names assigned to each cohort in the experiment. P1 and P2 should be used to describe the two parental strains whereas other names can be chosen by the user to best describe the crosses. The mean and SE columns describe the phenotype measured in each cohort. The sex column should contain "U," "E," "M" or "F" to indicate that the measured cohort had an unequal sex ratio, equal sex ratio, was composed of all males or all females, respectively. The sire and dam columns describe the row of the table to which the sire and dam belonged. For instance, the F1 cohort (shaded row) was produced using males from table row 2 (P2) and females from table row 1 (P1) and the phenotype was measured in an unequal number of males and females. This annotation is sufficient for the construction of an accurate C-matrix

Cohort	Mean	SE	Sex	env	Sire	Dam
P1	20.8	6.47	U	35	1	1
P2	24.5	8.98	U	35	2	2
F1	42.0	6.27	U	35	2	1
rF1	31.7	7.30	U	35	1	2
F2a	25.7	5.99	U	35	3	3
F2b	25.47	5.89	U	35	3	4
rF2a	38.42	4.60	U	35	4	3
rF2b	34.74	5.63	U	35	4	4

2.2 | C-matrix generation

The development of an appropriate C-matrix that represents the potential contribution of genetic effects on the mean phenotypes of cohorts is the first step in an LCA experiment. The accuracy and completeness of this matrix are central to performing LCA in an unbiased fashion. To eliminate errors in the construction of C-matrices, our software takes user input describing the crosses performed and uses algorithms to fill a C-matrix matching the user's data. Briefly, the user supplies the data that includes sire and dam identity for each cross (Table 1). Then, to produce a C-matrix from this data, we first fill in the rows representing the diverged parental strains (we refer to these throughout as P1 and P2). These two strains are assumed to be homozygous for alleles that cause divergence in the trait of interest. Though the assumption of homozygosity applies only to those loci that impact the trait of interest, this assumption may be violated in some experiments. Violations of this assumption have not been studied but are likely to cause a poor fit between the expected contribution of CGEs among cohorts and the measured phenotypes leading to reduced power. If we define P1 as the strain with the larger phenotype measure and P2 as the strain with the smaller phenotype measure, we can assign coefficients for additive effects as positive one for P1 and negative one for P2. Because these two cohorts are presumed homozygous at any loci of interest, we assign a coefficient of zero for dominance effects. Using this approach, we can also assign coefficients for cytotype, X chromosome, Y chromosome

and other effects (a full C-matrix showing 34 possible genetic effects (9 simple CGEs and 25 epistatic CGEs) for 16 typical cohorts is given in Table S1). To facilitate the calculation of subsequent cohorts, we also assign and track the proportion of the genome that originates from parental strain P2 (one for P2 and zero for P1). Likewise, a similar value is calculated and tracked for the X chromosomes and the Y chromosome. It is important to note that these values assigned for the X and Y chromosome are for the sex-limited portion of these chromosomes and not the pseudoautosomal region. Any phenotypic differences determined by loci in the pseudoautosomal region will be largely inferred as an autosomal effect if the genetic distance from the sex-determining locus is large. With the C-matrix rows for P1 and P2 filled as described above, we can begin filling other rows using a series of formulas.

To calculate the coefficient for the autosomal additive effect (A_a) we transform the proportion of the cohorts genome that originates from P2 onto the range (−1, 1) (4).

$$A_a = S_{P2} + D_{P2} - 1 \tag{4}$$

S_{P2} is the proportion of the sire's genome that comes from the P2 line whereas D_{P2} is the proportion of the dam's genome that comes from the P2 line. Our software applies a modified version of this formula in the case of sex chromosome effects that correctly accounts for the unequal inheritance of sex chromosomes among males and females. For instance, to calculate the X chromosome additive effect (X_a), we use (5) for females and (6) for males.

$$X_a = S_{XP2} + D_{XP2} - 1 \tag{5}$$

$$X_a = 2D_{XP2} - 1 \tag{6}$$

Similarly, S_{XP2} is the proportion of the sire's genome that comes from the P2 line whereas D_{XP2} is the proportion of the dam's genome that comes from the P2 line. In cases where a cohort is represented by an equal number of males and females, X_a is calculated as the mean of Equations (5) and (6).

The autosomal dominance (A_d) coefficient represents the probability that a randomly chosen site in the genome will be heterozygous for alleles from the two parental strains.

$$A_d = S_{P2}(1 - D_{P2}) + D_{P2}(1 - S_{P2}) \tag{7}$$

On the right side of (7), the first term provides the probability that the sire provides a P2 allele and the dam provides a P1 allele, whereas the second term represents the alternative, where the dam is providing the P2 allele and the sire is providing the P1 allele. This effectively represents the two ways that a heterozygote can occur. This equation can also be simplified to (8).

$$A_d = S_{P2} + D_{P2} - 2S_{P2}D_{P2} \tag{8}$$

For the case of X chromosome dominance effects in females, we can use this same formula but substitute the S_{XP2} for S_{P2} and D_{XP2} for D_{P2} yielding (9).

$$X_d = S_{XP2} + D_{XP2} - 2S_{XP2}D_{XP2} \tag{9}$$

Since it has no opportunity to contribute to the male phenotype, X_d is calculated as one half of Equation 9 in cohorts that contain equal numbers of males and females.

For Y chromosome, cytotype and mitochondrial effects, we carry down these uniparentally inherited values from the appropriate parents (under the assumption of no heteroplasmy).

One of the essential extensions that we provide to LCA is the ability of the user to supply multiple values for each cohort. For instance, the user could supply the mean phenotype value for both males and females separately. Alternatively, a user might raise genetically identical cohorts in different environmental conditions and provide each of these individually in their input data. We incorporate an environmental variable supplied by the user by rescaling the measured environmental value for each cohort on the interval (−1, 1) using (10).

$$v_i' = \frac{2\left(v_i - \min\left(v_i \ldots v_n\right)\right)}{\max\left(v_i \ldots v_n\right) - \min\left(v_i \ldots v_n\right)} - 1 \tag{10}$$

where $\max(v_i \ldots v_n)$ is the maximum value measured for the environmental variable and $\min(v_i \ldots v_n)$ is the minimum value measured for the environmental variable. In doing this, we can create a new column in the C-matrix that allows us to infer any simple environmental effects. Our approach assumes that the environmental variable is either binary discrete or continuous with an impact on phenotype that is linear. In cases where this is violated, the variable could first be transformed onto a scale where it behaves in a more linear fashion. With autosomal additive, autosomal dominance, X, Y, cytotype and environment coefficient effects calculated, all possible epistatic coefficients can also be filled in as the product of the single-locus or locus × environment (G × E) coefficients. Using this approach, we can fill in a complete C-matrix for any crossing design a user might choose—including all possible G × E CGEs.

As an example, we could study plant height with three different levels of watering: 15, 20 and 25 ml regimes. These would be rescaled as −1, 0 and 1, respectively, and assigned as coefficients in our C-matrix. The effect that we estimate for the environment with this approach would then be equal to the expected increase in plant height resulting from increasing watering by 5 ml. Additionally, we could assign interactions, such as an environmental by additive effect. This interaction term allows us to determine the degree to which the environmental variable and a composite genetic effect interact to determine the observed phenotype. Note that from the standpoint of a gene by environment interaction, we could also treat the sex of an organism identically (i.e. as

a binary environmental variable that can interact with the genes in any portion of the genome). By applying this approach to sex, we can investigate the genetic architecture of sexual dimorphism. If we measure a phenotype of interest in males and females of each cohort separately and assign a sex column that has negative one for one sex and positive one for the other, we can then recover the simple effect of sex as well as interactions with traditional genetic effects.

2.3 | Data and analysis

2.3.1 | Environmental variation

To evaluate the performance of our approach to incorporating environmental variation and gene by environment interactions, we created a variety of simulated data sets. The first simulated data set is generated under a model where the phenotype is described by a simple additive effect but is impacted by unrecognized environmental variation. To simulate data, we first chose the set of crosses that we would evaluate. We chose an experimental design with eight cohorts: parental strains (P1 and P2), initial crosses (F1 and rF1) and then four backcrosses involving an F1 or rF1 sire and a parental strain dam. To create the expected phenotypes for each of these cohorts, we assigned a midparent mean of 2.25 and an additive effect of 0.35. These values allow us to assign an expected mean value to any cohort; for instance, the P2 line would have a mean phenotype of 2.6, an F1 would have a mean phenotype of 2.25, and a backcross to the P1 parent would have a mean phenotype of 2.075. Next, we sampled 100 individuals from each cohort, drawing individual phenotype measures from a normal distribution with a mean equal to the expected value for the cohort and a standard deviation equal to 0.3. The above values (mean, additive effect and standard deviation) were informed by a recent analysis of ovule number in crosses between strains of Silene (Delph & Demuth, 2016). To incorporate environmental variation, we imagined an unmeasured binary environmental variable that impacted the phenotype of interest. We evaluated ten magnitudes of the environmental effect ranging from zero to 0.35 (the same as the genetic effect). At each of the ten levels, we created data sets where a proportion of randomly chosen individuals (regardless of cohort membership) experienced the environmental effect (had this value added to their sampled phenotype) or where all individuals of randomly chosen cohorts experienced the environmental effect (the expected value, i.e. mean for the normal distribution was increased by this value). Some environmental variables may have effects that vary both in magnitude and sign (e.g. some individuals have an increased phenotype whereas others have a decreased phenotype). Because this type of environmental variation increases cohort variance, it will reduce power but should not lead to biases in the inference of trait genetic architecture. Therefore, we have instead focused our analysis on environmental effects with a consistent directional impact on the trait of interest which has the potential to create strong biases in the inference of trait genetic architecture.

To explore the role that the frequency of the environmental effect has on our analysis, we repeated this process varying the proportion of individuals or cohorts that experienced the environmental effect. In the case of individuals experiencing the environmental effect, we created data sets where the environmental effect impacted between 0% and 100% of individuals. For the case of cohorts experiencing the environmental effect, we created data sets where the environmental effect impacted zero to eight randomly chosen cohorts. The cohort treatment was designed to capture the genuine concern that, in some experiments, later generation crosses may be temporally separated from early crosses leading to the opportunity for whole cohorts to experience an unmeasured environmental change. For each simulation condition, we generated 1000 data sets yielding a total of 180,000 simulated data sets.

To understand the impact of smaller sample sizes in the presence of unrecognized environmental variation, we also performed a smaller simulation study with cohort sizes of 10, 50 and 100 individuals. With each sample size, we followed a simulation procedure similar to that described above. However, the environmental effect was held constant 0.17 (approximately half of the true genetic effect). This environmental variable was allowed to impact the phenotype of three randomly chosen cohorts in each simulation. For each of these scenarios, we simulated 100 data sets and performed line-cross analysis on each simulated data set.

In some cases, environmental variation is simply unavoidable. To demonstrate the strength of our approach to incorporating known environmental variation, we re-evaluate classic data sets exploring the genetic architecture of fruit weight, number of locules per fruit, fruits per centimetre, plant spread and plant height in crosses between Johannisfeuer and Danmark tomato strains with each cross grown in two successive years (Powers, 1941). Each of these data sets is first analysed alone and then combined with the addition of only a binary environmental variable representing the year that the cohort was grown.

2.3.2 | Adaptive divergence

Next, we simulated data sets to show how our approach can be applied to understand the genetic architecture of adaptive divergence. We created a simulated data set inspired by investigations of local adaptation in sticklebacks (Rundle, 2002). Like this original investigation, we pictured a species where each parental strain is adapted to its typical environment, and we then measured some correlate of fitness (e.g. growth rate) of both strains and crosses between them in both environments. For this simulation, we used five cohorts: P1, P2, F1, BC1 and BC2. BC1 and BC2 are backcrosses of the F1 to P1 and P2, respectively. For each cohort, we generated the expected values by assigning a midparent mean of ten and an environment by autosomal additive effect of four. This environment by autosomal additive effect represents the case that some genes carried on autosomes have alleles with additive effect that benefits the organism but only in certain environments. We created our simulated data sets by using these expected values as the mean of a normal distribution with a standard deviation of six. For each cohort, we sampled 10, 25, 50 or 100 individuals. We chose these values based on data presented in

their Figure 2 (Rundle, 2002). For each sample size, we simulated 100 replicate data sets for a total of 400 simulated data sets.

2.3.3 | Sexual dimorphism

Finally, to illustrate the use of our approach in understanding the genetic architecture of sexual dimorphism, we created simulated data sets representing a species where males possess an exaggerated trait, but populations differ in the degree of dimorphism. For simplicity, we kept the magnitude of the phenotype and the cohorts sampled the same as described above for our investigation of unrecognized environmental variation. However, each cohort was now represented by separate measures for males and females. To create the expected phenotypes for each of these cohorts, we assigned a midparent mean of 2.25, and we also assigned a sex effect of 0.15 so that males would have a generally higher phenotype measure than comparable females. Next, we explored three ways that a genetic difference in the degree of sexual dimorphism might be generated. The first genetic architecture had a sex by autosomal additive effect. The second genetic architecture had a sex by X additive effect. The final genetic architecture was a Y effect which does not require a sex interaction since it is already sex-limited. In each of these cases, the effect size was set to 0.15. In all cases, individual phenotypes were drawn from a normal distribution with a mean equal to the expected value and a standard deviation of 0.3. Under this simulation, the difference in sexual dimorphism (between P1 and P2) is of roughly the same magnitude as the minimum difference in the sexes. For each of these three genetic architectures, we simulated data sets where 10, 25, 50 or 100 individuals were sampled in each cohort. For each set of simulation conditions, we simulated 100 replicate data sets for a total of 12,000 simulated data sets.

In all cases, resulting data sets were analysed with R version 3.4.4 and the LCA function in the package SAGA version 2.0 (https://github.com/coleoguy/SAGA2). For the first simulated data set, the generating model has just a single effect (autosomal additive), and we set the max.pars argument to 4. Briefly, the max.pars argument limits the number of possible variables included in models tested. This is helpful in cases where many cohorts make it possible to test many potential CGEs simultaneously leading to a model space consisting of hundreds of thousands of possible models. We tested a small number of data sets with larger max.pars arguments and saw no systematic differences in parameter estimates, but by limiting the size of models, we were able to analyse the 180,000 data sets in a reasonable time. Because there were far fewer data sets to analyse in other simulated and empirical examples, we did not have to use the max.pars argument in other analyses. For both the simulated data set investigating adaptation and the empirical datasets, we used the default settings except for the env argument that was set to true. For the purpose of calculating power and false-positive rate, we considered a CGE significant if it had a variable importance greater than 0.9 and a parameter estimate where the confidence interval excluded zero (Blackmon & Demuth, 2016). Scripts used for all simulations and analyses are found in the Dryad repository associated with this manuscript (https://doi.org/10.5061/dryad.fd0s01b).

3 | RESULTS

Our simulation approach included environmental variation both among individuals and among cohorts. Our results suggest that variation among cohorts has far greater potential to reduce power and create false positives than does environmental variation among individuals from all cohorts. For the case where undocumented variation that occurs across all individuals regardless of cohort membership, we find that false-positive rate ranges from 0.018 to 0.032 (Figure 1a; red lines). Furthermore, we find no consistent differences in weak and strong environmental effect when environmental variation is spread across all cohorts equally. Under this same simulation condition, we find that power ranges from 0.78 to 0.83 and again we find no consistent difference among simulations with weak or strong environmental effect (Figure 1b; red lines). In contrast, an undocumented environmental variation that occurs in all individuals of some cohorts is more concerning and biologically more relevant since many crossing experiments involve individuals reared at different times a fact that may lead to unavoidable environmental variation. The proportion of simulated data sets that result in at least one false-positive reaches a high of 0.74 when just a single family is impacted by a strong and undocumented environmental effect (effect size = 0.35; Figure 1a; blue). The false-positive rate reduces from this high but remains significantly above 5% under all simulations that have whole families impacted by undocumented environmental variation. As the number of families impacted by undocumented environmental variation increases, the power reduces down to a low of approximately 27% when the environmental effect is strong (0.35) and 50% of the families are affected (Figure 1b; blue). We also performed a supplemental analysis to determine the impact of smaller sample sizes. As expected power is lowest and false-positive rate is highest when sample size is small. With a cohort size of ten, the false-positive rate was 21% this reduced to 14% when cohort size was 100. In contrast, power was 32% when cohort size was ten but reached a high of 42% when cohort size was 100 (Figure S1).

To illustrate the power of incorporating environmental variation, we show how even accounting for unmeasured but inevitable environmental variation can improve LCA. In this case, we reanalysed classic data sets of crosses between Johannisfeuer and Danmark tomato strains that were repeated in two separate years (Powers, 1941). In our reanalysis of this data, we found that fruit weight and plant spread both increased in the second year and as expected we recover this environmental effect in our analysis (Table S2). Additionally, when we combine the separate years of crosses into a single analysis with an effect assigned to the year of the cross, we are able to increase our certainty in the identification of genetic effects and, more importantly, in some cases we are able to recover a G × E interaction. Specifically, in the cross examining fruit weight, we find that there is an autosome additive by environment interaction

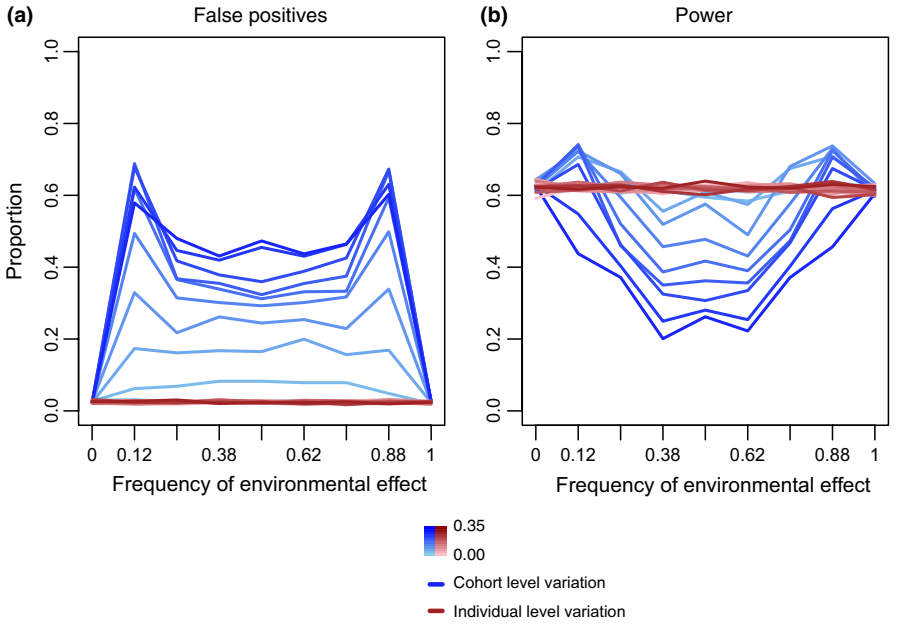

FIGURE 1 Impact of unrecognized environmental variation on LCA analyses as a function of the frequency and strength of environmental effects. Blue lines illustrate the result of whole cohorts experiencing an environmental effect. Red lines indicate the result of random individuals from all cohorts experiencing an environmental effect. The darkness of the red and blue lines indicates the strength of the environmental effect ranging from 0.0 to 0.35 (the same magnitude as the true genetic effect). (a) Proportion of simulated data sets that result in at least one false positive. (b) Proportion of simulated data sets where the true generating effect (additive autosomal) was inferred as important

(Figure 2). This suggests that some unmeasured environmental variable in these 2 years caused a greater shift in phenotypes measured in one strain than the other. For example, this would be expected if one strain had genes on autosomes that acted in an additive fashion to increase drought tolerance and one year was significantly drier than the other. This type of analysis also reveals phenotypic stability. For instance, the number of locules per fruit is different in the two strains (Johannisfeuer mean = 9.07 and Danmark mean = 5.83). Analysis of each year's data indicated a role for additive effects (1.85 and 1.43 in respective years) and dominance effects (−1.69 and −1.57 in respective years). However, when we combine these years, we do not see any sign of an effect of the environment on this trait and estimates for additive and dominance effects remain largely unchanged (1.57 and −1.61).

Including environmental variation can be even more powerful when the environment is intentionally manipulated. Using simulated data, we found that our LCA approach was able to successfully detect additive by environment interactions 86% of the time at biologically realistic effect sizes when sample size per cohort was at least 25 individuals (Figure 3; blue line). When sample size was lower than 25, we found that both power and false-positive rate suffered. When sample size per cohort was ten, 12% of simulated data sets resulted in false positives. At higher sample sizes, false-positive rates reduced to more acceptable levels. Despite relatively low power with small sample size, we found that our LCA approach was still effective at estimating the generating CGE effect. Specifically, when sample size per cohort was ten in those simulations where SAGA failed to find the generating CGE as significant, the mean parameter estimate was still 2.22 (standard deviation of 1.1), just slightly below the true value 3.0.

Our application of LCA to the analysis of sexual dimorphism illustrates the potential for LCA to uncover the genetic architecture of complex phenotypes that have traditionally been analysed in isolation. Our results indicate that in systems where sample size is sufficient (greater than ten per sex and cohort), we are able to reliably recover CGEs that describe the interaction between autosomal or sex chromosome loci with additive action and the sex they are carried in (Figure 4). When sample size per cohort is ten, we found that power to detect the genetic architectures was less than 50% (Autosome by sex = 34%, X chromosome by sex = 37%, Y chromosome = 48%). This quickly improved with increased sample size, with power in excess of 77% for all architectures when sample size per a cohort was 25 individuals. When sample size reached 50 individuals per cohort, all architectures were successfully identified in greater than 87% data sets. Although all three tested CGEs were largely similar, we found slightly higher power to detect sexual dimorphism that is due to genes that act in an additive fashion and occur on the Y chromosome in comparison with additive loci on autosomes or X chromosomes.

4 | DISCUSSION

Despite the early recognition that the mathematical framework of LCA could be extended to incorporate environmental effects (Bulmer, 1980), until now there has been no easy method for researchers to do so. In the new version of our software, we provide this extension and remove other barriers to LCA associated with construction of a matrix of genetic effects.

Our results illustrate that researchers must be cognizant of the potential for unrecognized environmental variation that impacts the trait of interest to influence the results of LCA. However, our results also show that if experiments are carefully designed to ensure that all cohorts are exposed to the same environmental variation, this is a danger that can largely be mitigated. In our simulation study, when environment was distributed randomly across individuals without respect to their cohort membership, we found no impact on estimates of CGEs. Although this is promising, it may be in part due to

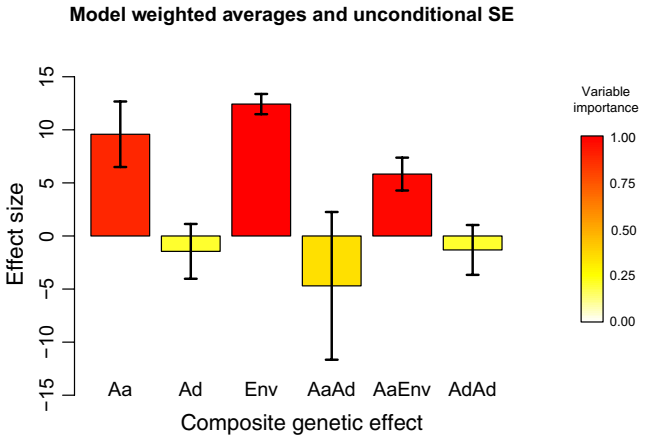

FIGURE 2 Composite genetic effects contributing to differences in fruit weight in crosses of Johannisfeuer and Danmark tomato strains. Parameter estimates are based on the a 95% confidence model set containing 69 of the 4478 tested models. Error bars indicate unconditional standard errors whereas colour of each bar indicates the variable importance . Composite genetic effects are indicated on the horizontal axis with abbreviations: A is autosomal, a is additive, d is dominance, Env is environment (e.g., AaEnv is autosomal additive by environment)

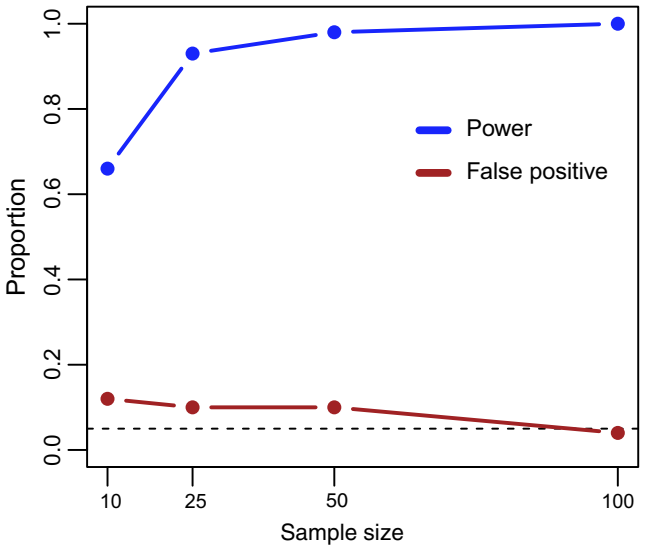

FIGURE 3 Inference of gene by environment interaction. The blue line illustrates the proportion of simulated datasets where an autosomal additive by environment interaction was detected as a function of the number of individuals measured in each cohort. The red line indicates the proportion of simulated data sets where any other composite genetic effect not included in the generating model was inferred to be important. The dashed black line indicates a value of 0.05

our sample size. We chose to simulate 100 individuals from each cohort in our primary analysis. Whereas this value is realistic for many study systems, lower sample sizes may be typical in some study systems. When cohort sample size is reduced from 100 to 10, we found that power fell by 10% and that the false-positive rate increased by 7%. These results underscore the importance of sufficient sample

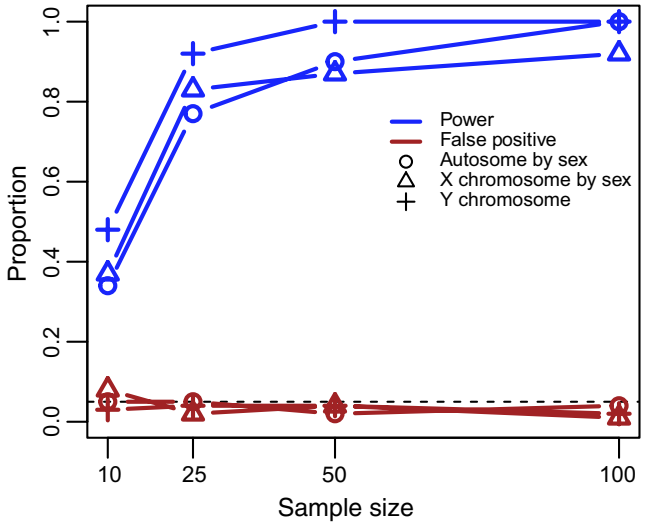

FIGURE 4 Inference of the genetic architecture of sexual dimorphism. The blue line illustrates the proportion of simulated data sets where the generating CGE was detected as a function of the number of individuals measured in each cohort. The red line indicates the proportion of simulated data sets where any other CGE not included in the generating model was inferred to be important. The shape of the symbol indicates the CGE that was included in the generating model. The dashed black line indicates a value of 0.05

size and controlling environmental variables that may impact the trait being studied when sample sizes are small.

Perhaps the most promising result of our investigation of unrecognized environmental variation comes from our reanalysis of empirical data sets. Here, we combined data from multiple years of crops grown outdoors—inevitably experiencing many differences in the environment. We found that by simply allowing year to be a variable in the analysis, we could infer the impact of changes in the environment across years, and in the case of fruit weight, we were even able to infer that there was an interaction between the genes of one line and the changes in the environment (Figure 2). We believe that this type of analysis where time is the environmental variable may be particularly helpful when conducting experiments where cohorts cannot be generated simultaneously and instead an experiment is spread across a number of months. In these cases, the date associated with each cohort can be used allowing an experimenter to control for unintended or unrecognized changes in the environment. However, the model as designed assumes a linear relationship between the environmental variable (e.g. date) and the phenotype being studied. In cases where this is unlikely to be true, it should not be applied or date could be transformed to a new scale where its impact on the phenotype would be expected to be linear. Furthermore, in the case where a researcher measures a possible confounding variable, this framework provides a simple and straightforward approach to test whether the possible confounding variable impacted the observed phenotypes of their cohorts. It should be noted that this approach to incorporating an environmental variable is suited to either a binary environmental variable allowing two levels of effect in the model or as a continuous variable that has a linear effect on the phenotype of interest.

One concern in the long running debate over the importance of epistasis has been the idea that some methods of studying genetic architecture might be biased towards or against the inference of complex epistatic architectures. With regard to LCA, there has been concern that noise in the form of unintended or unmeasured environmental variation could make the inference of epistatic genetic architectures more likely. To investigate this, we re-evaluated all of our simulation results of undocumented environmental variation to determine whether there was clear pattern in the types of genetic architectures that were inferred to be important despite their absence from the generating model. In the crossing design that we simulated, five nonepistatic and eleven epistatic CGEs can be distinguished. To describe these, we use a capital letter to indicate the portion of the genome and a subscript to indicate the mode of gene action. Epistatic interactions are denoted by combining two nonepistatic effects. These nonepistatic effects are autosomal additive (A_a) and dominance (A_d), X chromosome additive (X_a) and dominance (X_d), and cytotype additive (C_a); the 11 epistatic CGEs that can be distinguished are A_aA_a, A_aA_d, A_aX_a, A_aX_d, A_aC_a, A_dA_d, A_dX_a, A_dX_d and A_dC_a. Of these CGEs, only A_a was included in the generating model. Thus, there are four nonepistatic and eleven epistatic potential false positives (73.3% of possible false positives are epistatic). If we examine the case of strong environmental effect (described in detail above) and look across all levels of undocumented cohort level variation (Figure 1a; dark blue lines), we find that similar to their proportion in the overall set, 74% of false positives are for epistatic CGEs. This simple measure suggests that perhaps LCA is not biased towards inference of epistatic interactions. However, if we restrict our analysis to simulations where only one cohort is impacted by environmental variation, we find that the proportion of false positives that are epistatic rises to 89%. In contrast, when two cohorts are affected by environmental variation the proportion of false positive that are epistatic falls to 51%. Although we cannot completely explain this phenomenon, it seems to result from the fact that certain epistatic CGEs are strongly differentiated by just one or a handful of cohorts and if one of these cohorts is impacted by environmental variation it can drive a significant result. The specific CGEs and cohorts involved will change depending on the experimental design, but the potential severity of this problem should be lessened by larger experiments that have more cohorts differing in the contribution of each CGE.

Despite sharing the majority of the genome, sexual dimorphism is common in almost all clades with separate sexes (Barrett & Hough, 2012; Darwin, 1871). Theoretical predictions suggest the loci underlying sexually dimorphic traits might be concentrated on the sex chromosomes, since this can reduce negative intersexual correlations in fitness (Blackmon & Brandvain, 2017; Charlesworth & Charlesworth, 1980; Fisher, 1958; Rice, 1984; Rice & Chippindale, 2001). Unfortunately, most theoretical work has focused on the evolutionary dynamics of single loci, and it remains unclear how the polygenic nature of most traits might change these predictions. Furthermore, empirical evidence has shown that autosomes are often important contributors to sexual dimorphic traits, and even species without sex chromosomes exhibit sexual dimorphism (Mank, 2008). The extension we have developed for LCA allows a novel approach to investigate the genetic architecture of sexual dimorphism that should shed light on the relative contribution of autosomes and sex chromosomes even in the case of highly polygenic traits. In traditional LCA analyses, it has been necessary to either analyse only females, only males, or cohorts made up of an equal number of males and females. Whereas these approaches are all still possible under our implementation, the ability to combine analysis of separate male and female means for each cohort in a single analysis should be the preferred approach used moving forward. This method will allow for a better understanding of the interactions between sex and genetic architecture and a more accurate estimation of many genetic effects.

Understanding interactions between genomes and the environment that produce the selective forces that act on them are central to evolutionary biology. We have illustrated that with LCA, we can successfully recover the signal of this relationship with biologically realistic effect sizes. This is particularly important since it offers an opportunity to explore intrinsic versus extrinsic reproductive isolation and the process of adaptive speciation. Furthermore, because the LCA approach looks at the average effects across all loci, it is not limited to finding loci of large effect; instead, we can see the contribution of all loci across the genome simultaneously. This offers a robust approach to detecting complex interactions that might be missed with methods that search for QTLs or QTNs. Furthermore, it is a method available to any researcher without the necessity of sequencing many individuals, an important consideration despite the falling prices of sequencing. In sum, environmental variation does introduce some peril to an LCA study, but if handled properly the incorporation of environmental data can be dealt with and can even increase the inferential power of a study. Furthermore, the same mathematics that allows us to deal with environmental variation also allows us to leverage LCA to characterize the genetic architecture of sexual dimorphism and adaptive divergence in any study system where controlled crosses can be produced and phenotyped.

ACKNOWLEDGMENTS

This work benefited from and was motivated by discussions with Ruth Shaw about the challenges and possible problems of applying line cross analysis in the presence of environmental variation. Jeffery P. Demuth and two anonymous reviewers provided comments and suggestions that greatly improved the manuscript.

ORCID

Heath Blackmon https://orcid.org/0000-0002-5433-4036

REFERENCES

Barrett, S. C., & Hough, J. (2012). Sexual dimorphism in flowering plants. *Journal of Experimental Botany*, *64*, 67–82.

Blackmon, H., & Brandvain, Y. (2017). Long-term fragility of Y chromosomes is dominated by short-term resolution of sexual

antagonism.*Genetics*,genetics.300382.2017.https://doi.org/10.1534/genetics.117.300382

Blackmon, H., & Demuth, J. P. (2016). SAGA: software for the analysis of genetic architecture. *Evolution*, *70*, 420–432. https://doi.org/10.1111/evo.12844

Brodie, E. D., Wade, M. J., & Wolf, J. B. (2000). *Epistasis and the evolutionary process*. Oxford: Oxford Univ. Press.

Bulmer, M. G. (1980). *The mathematical theory of quantitative genetics*. Oxford: Clarendon Press.

Cavalli, L. (1952). An analysis of linkage in quantitative inheritance.

Charlesworth, D., & Charlesworth, B. (1980). Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes. *Genetical Research*, *35*, 205–214. https://doi.org/10.1017/S0016672300014051

Cheverud, J. M., & Routman, E. J. (1996). Epistasis as a source of increased additive genetic variance at population bottlenecks. *Evolution*, *50*, 1042–1051. https://doi.org/10.1111/j.1558-5646.1996.tb02345.x

Coyne, J. A., Barton, N. H., & Turelli, M. (1997). Perspective: a critique of Sewall Wright's shifting balance theory of evolution. *Evolution*, *51*, 643–671. https://doi.org/10.1111/j.1558-5646.1997.tb03650.x

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: Murray.

Delph, L. F., & Demuth, J. P. (2016). Haldane's rule: genetic bases and their empirical support. *Journal of Heredity*, *107*(5), 383–391. https://doi.org/10.1093/jhered/esw026

Demuth, J. P., Flanagan, R. J., & Delph, L. F. (2014). Genetic architecture of isolation between two species of Silene with sex chromosomes and Haldane's rule. *Evolution*, *68*, 332–342. https://doi.org/10.1111/evo.12269

Demuth, J. P., & Wade, M. J. (2006). Experimental methods for measuring gene interactions. *Annual Review of Ecology Evolution and Systematics*, *37*, 289–316. https://doi.org/10.1146/annurev.ecolsys.37.091305.110124

Demuth, J. P., & Wade, M. J. (2007). Population differentiation in the beetle *Tribolium castaneum*. I. Genetic architecture. *Evolution*, *61*, 494–509. https://doi.org/10.1111/j.1558-5646.2007.00048.x

Donnelly, P. (2008). Progress and challenges in genome-wide association studies in humans. *Nature*, *456*, 728. https://doi.org/10.1038/nature07631

Egan, S. P., & Funk, D. J. (2009). Ecologically dependent postmating isolation between sympatric host forms of *Neochlamisus bebbianae* leaf beetles. *Proceedings of the National Academy of Sciences, USA*, *106*, 19426–19431. https://doi.org/10.1073/pnas.0909424106

Falconer, D., & Mackay, T. (1989). *Introduction to quantitative genetics* (pp. 313–320). New York: John Willey and Sons. Inc..

Fisher, R. A. (1958). *The genetical theory of natural selection*, 2nd ed.. New York: Dover Publications.

Fuchsberger, C., Flannick, J., Teslovich, T. M., Mahajan, A., Agarwala, V., Gaulton, K. J., … McCarthy, D. J. (2016). The genetic architecture of type 2 diabetes. *Nature*, *536*, 41. https://doi.org/10.1038/nature18642

Gall, G. (1975). Genetics of reproduction in domesticated rainbow trout. *Journal of Animal Science*, *40*, 19–28. https://doi.org/10.2527/jas1975.40119x

Goodnight, C. J. (1988). Epistasis and the effect of founder events on the additive genetic variance. *Evolution*, *42*, 441–454. https://doi.org/10.1111/j.1558-5646.1988.tb04151.x

Hayman, B. (1958). The separation of epistatic from additive and dominance variation in generation means. *Heredity*, *12*, 371. https://doi.org/10.1038/hdy.1958.36

van Heerwaarden, B., & Sgrò, C. M. (2017). The quantitative genetic basis of clinal divergence in phenotypic plasticity. *Evolution*, *71*, 2618–2633. https://doi.org/10.1111/evo.13342

Kitano, J., Ross, J. A., Mori, S., Kume, M., Jones, F. C., Chan, Y. F., … Myers, R. M. (2009). A role for a neo-sex chromosome in stickleback speciation. *Nature*, *461*, 1079–1083. https://doi.org/10.1038/nature08441

Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, *9*, 29. https://doi.org/10.1186/1746-4811-9-29

Küpper, C., Stocks, M., Risse, J. E., dos Remedios, N., Farrell, L. L., McRae, S. B., … Verkuil, Y. I. (2016). A supergene determines highly divergent male reproductive morphs in the ruff. *Nature Genetics*, *48*, 79. https://doi.org/10.1038/ng.3443

Ledón-Rettig, C., Zattara, E., & Moczek, A. (2017). Asymmetric interactions between doublesex and tissue-and sex-specific target genes mediate sexual dimorphism in beetles. *Nature Communications*, *8*, 14593. https://doi.org/10.1038/ncomms14593

Mank, J. E. (2008). Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. *American Naturalist*, *173*, 141–150.

Mather, K., & Jinks, J. L. (1982) Diallels. In *Biometrical genetics* (pp. 255–291). London: Chapman and Hall.

Miller, G. T., Starmer, W. T., & Pitnick, S. (2003). Quantitative genetic analysis of among-population variation in sperm and female sperm-storage organ length in Drosophila mojavensis. *Genetics Research*, *81*, 213–220. https://doi.org/10.1017/S0016672303006190

Mundry, R., & Nunn, C. L. (2008). Stepwise model fitting and statistical inference: turning noise into signal pollution. *American Naturalist*, *173*, 119–123.

Ober, C., Loisel, D. A., & Gilad, Y. (2008). Sex-specific genetic architecture of human disease. *Nature Reviews Genetics*, *9*, 911. https://doi.org/10.1038/nrg2415

Pandey, A., Davis, N., White, B., Pajewski, N., Savitz, J., Drevets, W., & McKinney, B. (2012). Epistasis network centrality analysis yields pathway replication across two GWAS cohorts for bipolar disorder. *Translational Psychiatry*, *2*, e154. https://doi.org/10.1038/tp.2012.80

Peck, S. L., Ellner, S. P., & Gould, F. (1998). A spatially explicit stochastic model demonstrates the feasibility of Wright's shifting balance theory. *Evolution*, *52*, 1834–1839. https://doi.org/10.1111/j.1558-5646.1998.tb02260.x

Poissant, J., Wilson, A. J., & Coltman, D. W. (2010). Sex-specific genetic variance and the evolution of sexual dimorphism: a systematic review of cross-sex genetic correlations. *Evolution*, *64*, 97–107. https://doi.org/10.1111/j.1558-5646.2009.00793.x

Pooni, H., Jinks, J., & de Toledo, J. (1985). Predicting and observing the properties of second cycle hybrids using basic generations and inbred line× F 1 crosses. *Heredity*, *54*, 121. https://doi.org/10.1038/hdy.1985.16

Powers, L. (1941). Inheritance of quantitative characters in crosses involving two species of Lycopersicon. *Journal of Agricultural Research*, *63*, 149–174.

Rice, W. R. (1984). Sex chromosomes and the evolution of sexual dimorphism. *Evolution*, *38*, 735–742. https://doi.org/10.1111/j.1558-5646.1984.tb00346.x

Rice, W., & Chippindale, A. (2001). Intersexual ontogenetic conflict. *Journal of Evolutionary Biology*, *14*, 685–693.

Rockman, M. V. (2012). The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution: International Journal of Organic Evolution*, *66*, 1–17. https://doi.org/10.1111/j.1558-5646.2011.01486.x

Roff, D. A., & Emerson, K. (2006). Epistasis and dominance: evidence for differential effects in life-history versus morphological traits. *Evolution*, *60*, 1981–1990. https://doi.org/10.1111/j.0014-3820.2006.tb01836.x

Rundle, H. D. (2002). A test of ecologically dependent postmating isolation between sympatric sticklebacks. *Evolution*, *56*, 322–329. https://doi.org/10.1111/j.0014-3820.2002.tb01342.x

Rundle, H. D., & Whitlock, M. C. (2001). A genetic interpretation of ecologically dependent isolation. *Evolution*, *55*, 198–201. https://doi.org/10.1111/j.0014-3820.2001.tb01284.x

Singh, G., Bhullar, G., & Gill, K. (1986). Genetic control of grain yield and its related traits in bread wheat. *Theoretical and Applied Genetics*, *72*, 536–540. https://doi.org/10.1007/BF00289537

Turelli, M., & Barton, N. (2006). Will population bottlenecks and multilocus epistasis increase additive genetic variance? *Evolution*, *60*, 1763–1776. https://doi.org/10.1111/j.0014-3820.2006.tb00521.x

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 years of GWAS discovery: biology, function, and translation. *American Journal of Human Genetics*, *101*, 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005

Wade, M. J., & Goodnight, C. J. (1998). Perspective: the theories of Fisher and Wright in the context of metapopulations: when nature does many small experiments. *Evolution*, *52*, 1537–1553. https://doi.org/10.1111/j.1558-5646.1998.tb02235.x

Wei, W.-H., Hemani, G., & Haley, C. S. (2014). Detecting epistasis in human complex traits. *Nature Reviews Genetics*, *15*, 722. https://doi.org/10.1038/nrg3747

Weiss, L. A., Pan, L., Abney, M., & Ober, C. (2006). The sex-specific genetic architecture of quantitative traits in humans. *Nature Genetics*, *38*, 218. https://doi.org/10.1038/ng1726

Wolak, M. E. (2013). *The quantitative genetics of sexual differences: new methodologies and an empirical investigation of sex-linked, sex-specific, non-additive, and epigenetic effects*. Riverside: UC Riverside.

Wolf, J. B., Brodie, E. D. III, Cheverud, J. M., Moore, A. J., & Wade, M. J. (1998). Evolutionary consequences of indirect genetic effects. *Trends in Ecology & Evolution*, *13*, 64–69. https://doi.org/10.1016/S0169-5347(97)01233-0

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, *16*, 97.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Armstrong A, Anderson NW, Blackmon H. Inferring the potentially complex genetic architectures of adaptation, sexual dimorphism and genotype by environment interactions by partitioning of mean phenotypes. *J Evol Biol*. 2019;00:1–11. https://doi.org/10.1111/jeb.13421