Optimizing breeding schemes

Manual

Heritability: meaning and computation

- This manual explains the concept of heritability and shows how to compute it by different methods along with recommendations.
- Published on 22/June/2021

  Excellenceinbreeding.org/toolbox/tools/ Heritability

Excellence in Breeding PLATFORM

ExcellenceinBreeding.org

Heritability: meaning and computation

Author.

Giovanny E. Covarrubias-Pazaran // g.covarrubiasl@cgiar.org

Breeding Optimization Lead, CGIAR Excellence in Breeding Platform (EiB)

Editors.

Valentin Wimmer // KWS

Emily Ziemke // Corteva

Johannes Martini // The International Maize & Wheat Improvement Center (CIMMYT)

Sam Storr // EiB

BILL & MELINDA GATES foundation

CIMMYT MR
International Maize and Wheat Improvement Center

CGIAR
Excellence in Breeding Platform

Contents

Introduction 1

Definitions and interpretations of heritability 2

Misconceptions of heritability 5

Methods to measure heritability 8

Conclusion and recommendations 18

Bibliography 19

Glossary 21

Introduction

In plant breeding programs, cultivars and materials of interest are often grown and tested at multiple locations across several years. Such a series of trials is called a multi-environment trial (MET), where a year–location combination is referred to as an *environment*. To quantify and eventually compare the precision of METs, plant breeders often calculate narrow-sense heritability (h^2) or broad-sense heritability (H^2) on a genotype-mean basis. The latter is defined as the proportion of phenotypic variance that is attributable to an overall variance for the genotype, thus including additive, dominance, and epistatic variance (Holland et al., 2003; Falconer and Mackay, 2005; Schmidt et al., 2019). As a key factor in achieving high rates of genetic gain, enabling the timely development and release of varieties that meet consumer and farmer needs, a clear understanding of heritability is necessary in public sector breeding programs.

This manual has three purposes:

1. Provide clarity on the meaning of heritability.
2. Show how to calculate heritability using suitable methods that allow for common understanding and transparency.
3. Provide recommendations on robust methods for quantifying and comparing the precision of field trials in public sector breeding.

1. Definitions and interpretations of heritability

Multiple definitions of heritability exist, e.g., “the portion of the observed variance for which differences in heredity are responsible” (Knight, 1948), or “the extent to which a phenotype is genetically determined” (Lourenço et al., 2017). Moreover, there are several interpretations associated with heritability: (i) it is equivalent to the coefficient of determination of a linear regression of the unobservable genotypic value on the observed phenotype, (ii) it is also the squared correlation between predicted phenotypic value and genotypic value, and (iii) it represents the proportion of the selection differential (S) that can be realized as the response to selection (R) (Falconer and Mackay, 2005), among others (Schmidt et al., 2019). Although many definitions, interpretations and methods exist, all converge on the idea of quantifying the genetic signal from phenotype measurements (Figure 1).

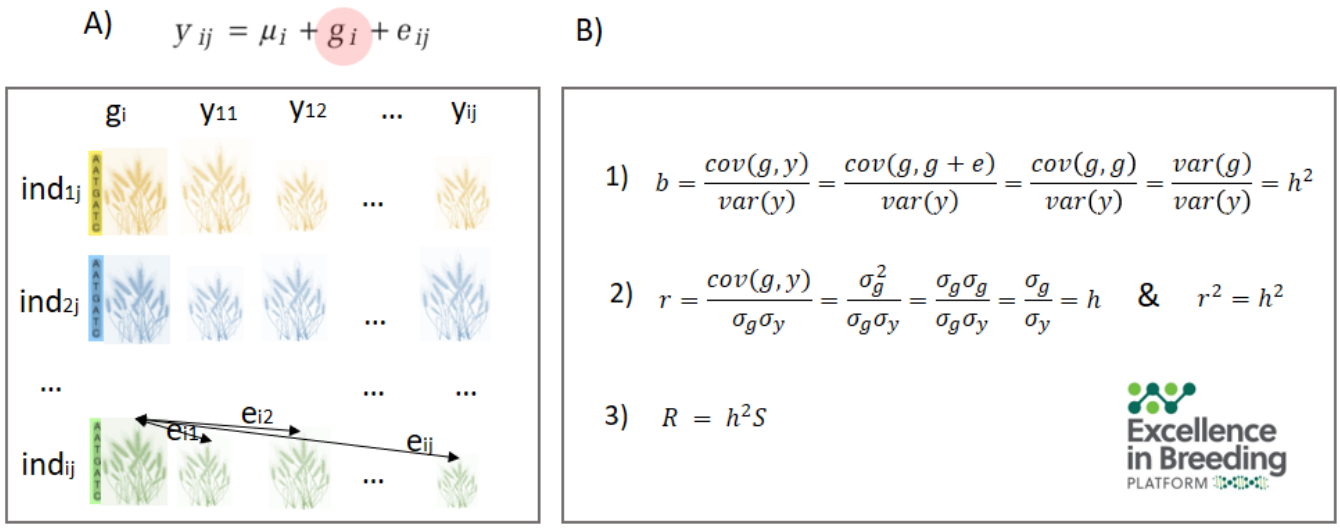

Figure 1. Graphical representation of phenotypic partition and three different heritability interpretations. In A) the phenotype (y_{ij}) is explained as the sum of an intercept (μ; mean) plus the effect attributed to the i^{th} genotype (g_i) plus the non-genotype effect attributed to other influences (e_{ij}) that confounds the genotype effect from other effects resulting in an observation. In B) 1) the heritability is described as the regression of the phenotype on the genotype, in 2) as the squared correlation between the phenotype and genotype and in 3) as the proportion of the selection differential that can be realized as the response to selection. All interpretations converge on the idea of quantifying the genetic signal from a phenotype.

The phenotypic variance in broad terms can be divided between genetic variance (the portion of the phenotypic variance attributed to genetic differences) and error variance (the portion of the variance that cannot be attributed to genetic differences but to other factors such as environment, etc.). Some methods to estimate heritability use the variance component for the plot error (σ^2_e) divided by the number of plots of each genotype to quantify the genetic signal, other methods use the average standard error of genetic estimates to derive the variance that cannot be attributed to genetic differences, and others use the slope of a regression (Figure 2).

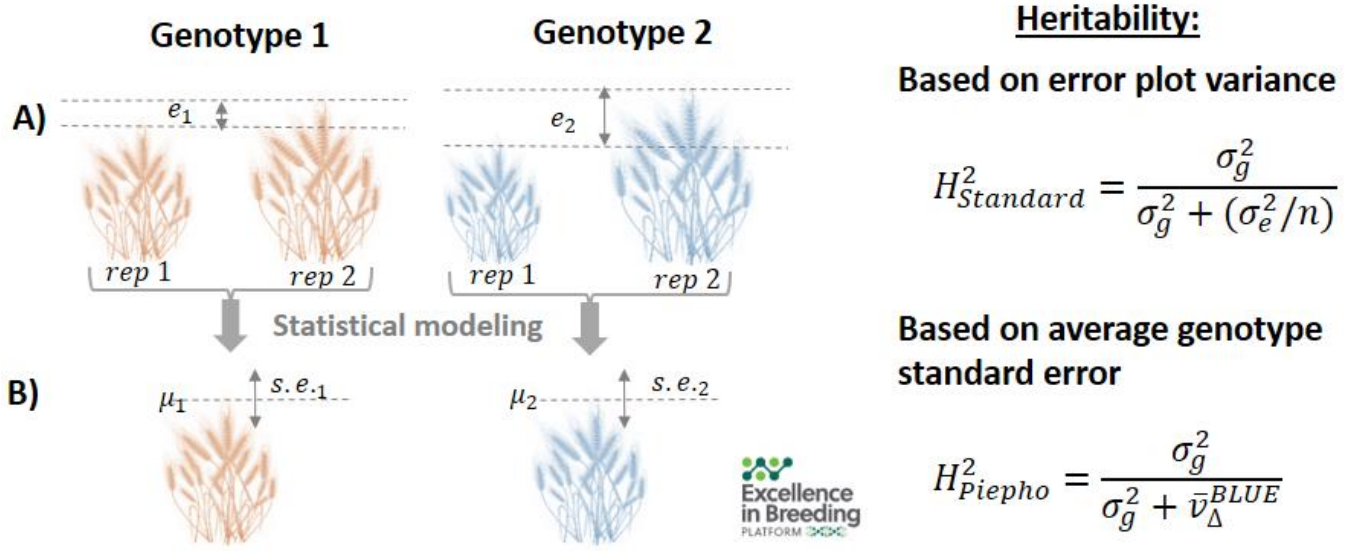

Figure 2. Example of two different ways to partition the genetic and the non-genetic variance needed for the computation of heritability. In A) the error variance component [i.e. estimated by restricted maximum likelihood (REML) or expected mean squares] is used in the denominator (σ^2_e) to quantify genetic signal, whereas in B) the standard errors ($s.e._i$) squared (variances) of the genetic estimates (μ_i; in the example BLUEs) after statistical modeling can be averaged to quantify the non-genetic variance and put in the denominator ($\bar{v}_\Delta^{BLUE}$) to quantify the genetic signal.

2. Misconceptions of heritability

Oldenbroek and van de Waaij (2015) summarize five major misconceptions regarding heritability:

Misconception 1. "A heritability of x indicates that x% of the trait is determined by genetics"

This is a very common misconception that arises from a misunderstanding of the definition of heritability. A heritability of 0.40 indicates that 40% of all the phenotypic variation for that trait is due to variation in genotypes for that trait. This differs importantly from the misconceived understanding that in each plant 40% of the expression of the trait is due to genes and the rest due to other influences.

Misconception 2. "A low heritability means that traits are not determined by genes"

A heritability that is larger than 0 always indicates that genes have an effect on the expression of the phenotype. The heritability is determined by the proportion of genetic variance relative to the phenotypic variance. A low heritability therefore can indicate that the genetic variance is low compared to the phenotypic variance (both could be small). For example, branching in maize is very much genetically determined, but because by far most genotypes used in modern maize programs have a single stem, the genetic variance for branching is very low.

Misconception 3. "A low heritability means that genetic differences are small"

A low heritability does not automatically indicate that the genetic variance is small; it can also indicate that the error variance is large. This can be caused by high environmental influence, for example, but also by inaccurate phenotype recording. For example, resistance to a certain infection will depend on the genetic potential to withstand that infection; the problem is how to measure that potential. If a single field measurement is taken of nematode infection in beat plants, it will record only those infected at that time, but this could vary according to the environment selected for recording infection levels.

Misconception 4. "A heritability is a fixed value"

The heritability reflects the relative weight of the genetic variance component in the phenotypic variance of a specific population and is based on observations that were taken on a specific moment in time. The magnitude of heritability depends on genetic variance in a population, but also on the influence of the environment and on the accuracy of observations (see misconception 3). The genetic variance in one population may be (somewhat) different from that in another population. Finally, heritability within a population can change over time, and for that reason, should be estimated at regular intervals.

Misconception 5. "A high heritability implies a major-effect QTL"

The fact that the heritability quantifies the genetic signal from a phenotype doesn't mean that says something about the genetic inheritance of the trait. Whether there's one or many thousands of genes behind and irrespectively of their effect we can have high or low heritabilities. A major-QTL trait like eye color can have low heritability if the population scanned have only one type of eye color, or a high heritability of we observe all types of color. A highly quantitative trait like yield can have a high heritability is the experiment is well conducted with high appropriate replication levels, but can also have low heritability if the agronomic management is poor.

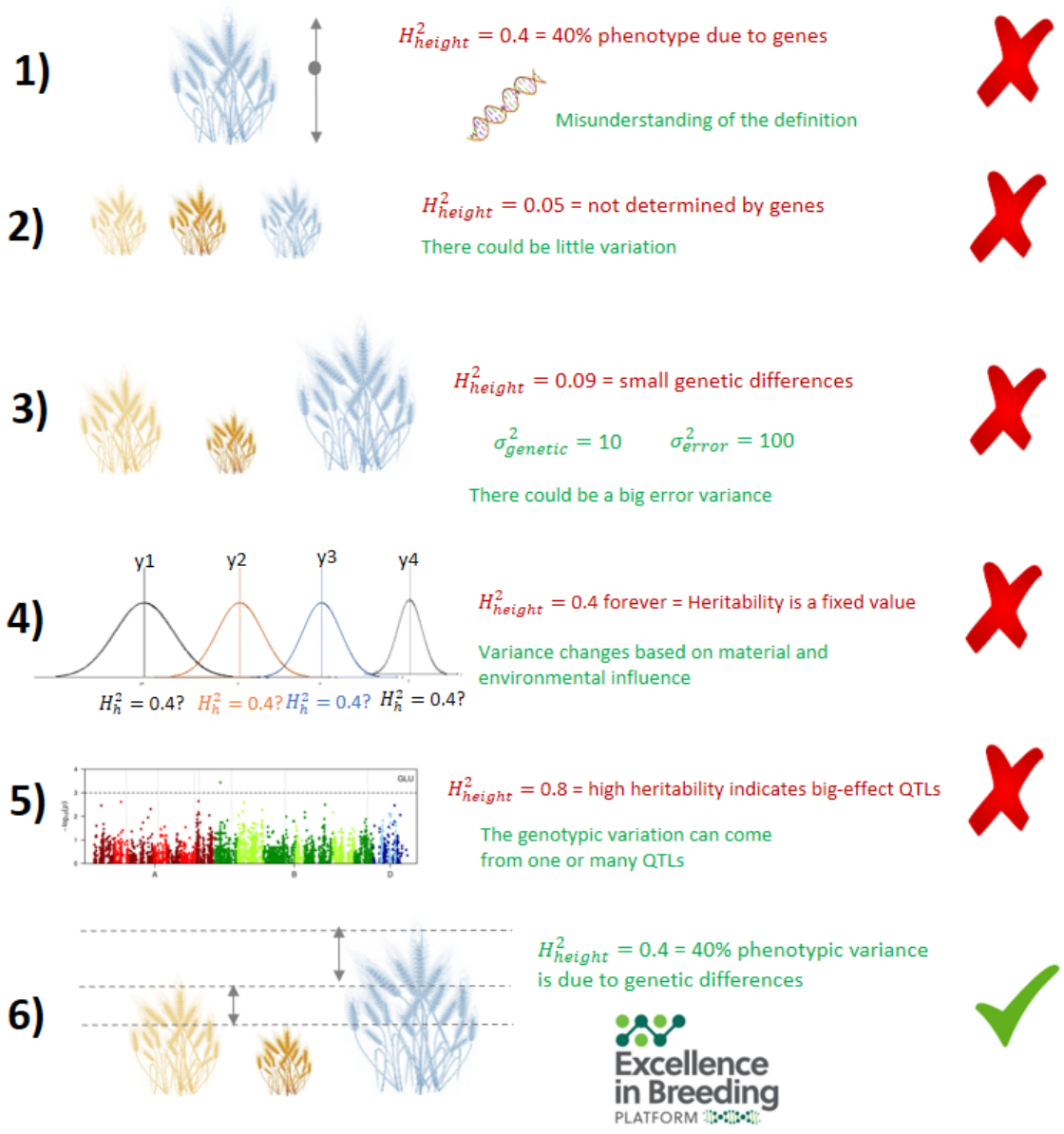

Figure 3. **Common misconceptions of heritability.** In 1), a misunderstanding of the concept results in the conclusion that a percentage of the phenotype is due to genes. In 2), the lack of variation resulting in low heritability is misunderstood to be consequence of no genetics contributing to the expression of the trait. In 3), a low heritability is misunderstood to reflect a small difference between genotypes when it could also be attributed to a large error variance. In 4), heritability is wrongly thought to be the always the same across time or populations. In 5), heritability is wrongly interpreted to be correlated to the number of large QTLs. In 6), the correct interpretation of heritability is provided.

3. Methods to measure heritability

Heritability is a useful concept in plant breeding and genetics, but given the many ways to generate phenotypic data (e.g. multiple reps, multiple years, multiple locations, along with the different levels of balance when measuring this metric) it has become difficult for breeders to choose which method to use. Here, we summarize some of the most suitable methods, of which the first is recommended as the most robust. Each method is demonstrated with an example, with a dataset that varies across examples in response to different issues with each method. Please keep in mind that heritability can be calculated for a single trial or for a multi-trial (or multi-environment; MET) experiment. We will touch on the single trial but the same principles can be extended to the MET scenario where the breeder will have to decide if heritability estimate has to be for each trial or for an across-trial. An excellent review of methods and code can also be found in Schmidt et al. (2019).

Method 1. "Cullis'" broad-sense heritability method (Recommended)

Cullis et al. (2006) propose a modern method that is widely used to account for the unbalanced scenario that plant breeders face in the single and multi-environment context with the advantage of not requiring every entry to have one rep. In this case, the genetic term is fitted as a random effect (BLUP). It uses the square of the standard error of the genetic estimates (across environments) to attain an approximation of the non-genetic variation (Figure 2). The formula is as follows:

$$H^2_{Cullis} = 1 - \frac{\bar{v}_{\Delta}^{BLUP}}{2\,\sigma_g^2} = 1 - \frac{\overline{PEV}}{\sigma_g^2}$$

Where: "σ2" refers to variance, "g" to genotype, $\bar{v}_{\Delta}^{BLUP}$ to the average BLUP difference or pairwise prediction error variance, and $\overline{PEV}$ refers to the average prediction error variance from genotypes. Where the proof is as follows (Aparicio J. personal communication):

$$\bar{v}_{\Delta}^{BLUP} = Var(\hat{Y}_i - \hat{Y}_{i'}) = 2Var(\hat{Y}_i) = 2PEV_i$$

$$H^2_{Cullis} = 1 - \frac{\bar{v}_{\Delta}^{BLUP}}{2\sigma_g^2}$$

$$H^2_{Cullis} = 1 - \frac{2\overline{PEV}}{2\sigma_g^2} = 1 - \frac{\overline{PEV}}{\sigma_g^2}$$

The advantage of this formula is that it can be calculated as part of a regular field trial analysis, can deal with unbalanced datasets and takes advantage from the random variable properties such as estimation of variance components.

Example: Assume a multi-environment trial (MET) run in four years across three locations in each year, for two repetitions in each year-location combination. After running a mixed model with genotypes fitted as random effects and the rest of terms as desired (see data and script available in additional material) the following parameters are obtained:

genotype	BLUP+ mu	$(s.e.)^2$ | PEV
1	5.01	0.0321
2	4.52	0.0321
3	3.73	0.0321
4	4.52	0.0321
5	4.99	0.0321
$\frac{\sum_{i=1}^{n} s.e^2}{n}$	-	0.0321

σ^2_g= 0.142; σ^2_y= 0.4; σ^2_l= 0.2; σ^2_{yl}= 0.6; σ^2_{error}= 0.08; $\overline{PEV} = 0.0321$

$$H^2_{Cullis} = 1 - \frac{\bar{v}_{\Delta}^{BLUP}}{2*\sigma_g^2} = 1 - \frac{\overline{PEV}}{\sigma_g^2} = 1 - \frac{0.0321}{0.142} = 0.775$$

As seen here, the calculation is straightforward, but it requires each individual to have more than one repetition in order to estimate a standard error for the estimates. It can be deduced that the 43% of the phenotypic variation can be attributed to genetic differences, and that there will be a response to selection.

Method 2. "Standard" broad-sense heritability method

This method is by far the most commonly used in the plant breeding community. This method provides the advantage of a straightforward calculation along with an intuitive parameter selection. Disadvantages include a tendency to overestimate values when data is unbalanced (different number of reps per genotype in single or multi-environment setting), as it assumes balanced datasets.

$$H^2_{Standard} = \frac{\sigma^2_g}{\sigma^2_p}$$

With: $\sigma^2_p = \sigma^2_g + \frac{\sigma^2_{ga}}{n_a} + \frac{\sigma^2_{gl}}{n_l} + \frac{\sigma^2_{gal}}{n_a n_l} + \frac{\sigma^2_{plot}}{n_a n_l n_r}$

Where: "σ^2" refers to variance, "n" to number of, "g" to genotype, "a" to years, "l" to locations, and "plot" to the plot error.

Example: Assume a MET run over four years across three locations in each year, for two repetitions in each year-location combination. After running a mixed model (see data and script available in additional material) the following parameters are obtained:

σ^2_g= 50; σ^2_{ga}= 200; σ^2_{gl}= 100; σ^2_{gal}= 10; σ^2_{plot}= 300; n_a=4; n_l=3; n_r=2;

$$\sigma^2_p = \sigma^2_g + \frac{\sigma^2_{ga}}{n_a} + \frac{\sigma^2_{gl}}{n_l} + \frac{\sigma^2_{gal}}{n_a n_l} + \frac{\sigma^2_{plot}}{n_a n_l n_r} = 50 + \frac{200}{4} + \frac{100}{3} + \frac{10}{4*3} + \frac{300}{4*3*2} = 146.66$$

$$H^2_{Standard} = \frac{50}{146.66} = 0.34$$

As seen here, the calculation is straightforward, but it assumes the same number of repetitions per year, and the same number of locations per year. It can be deduced that 34% of phenotypic variation (49 out of 146.66 trait units) can be attributed to genetic differences.

Method 3. "Ad hoc Holland" broad-sense heritability method

Another method to calculate heritability was proposed by Holland et al. (2003) to address the possible differences in the number of replications or locations by year. The idea behind is to come up with a harmonic mean value for the denominators in the heritability formula.

$$H^2_{Holland} = \frac{\sigma^2_g}{\sigma^2_{p_n}}$$

With: $\boldsymbol{\sigma^2_{p_n} = \sigma^2_g + \frac{\sigma^2_{ga}}{\tilde{n}_a} + \frac{\sigma^2_{gl}}{\tilde{n}_l} + \frac{\sigma^2_{gal}}{\tilde{n}_{al}} + \frac{\sigma^2_{plot}}{\tilde{n}_{alr}}}$

And: $\boldsymbol{\tilde{n}_a = \frac{n_g}{\sum_{i=1}^{n} \frac{1}{n_{a_i}}}, \tilde{n}_l = \frac{n_g}{\sum_{i=1}^{n} \frac{1}{n_{l_i}}}, \tilde{n}_{al} = \frac{n_g}{\sum_{i=1}^{n} \frac{1}{n_{al_i}}}, \tilde{n}_{alr} = \frac{n_g}{\sum_{i=1}^{n} \frac{1}{n_{alr_i}}}}$

Where: "σ^2" refers to variance, "$\tilde{n}$" harmonic mean of, "g" to genotype, "a" to years, "l" to locations, and "plot" to the plot error.

Example: Assume a MET run over four years (n_a) across 1-3 locations (n_l) in each year depending on the year, for 1-3 reps (n_r) per genotype in each year-location combination, depending on the year-location combination. After running a mixed model (see data and script available in additional material) the following data and parameters are obtained:

genotype	n_a	n_l	n_r	n_{al}	n_{alr}
1	4	3	2	4*3=12	4*3*2=24
2	4	3	2	4*3=12	4*3*2=24
3	4	2	1	4*2=8	4*2*1=8
4	4	1	2	4*1=4	4*1*2=8
5	4	2	3	4*2=8	4*2*3=24
$\sum_{i=1}^{n} \frac{1}{n_{x_i}}$	1.25	2.66	2.83	0.66	0.375
n_g	5	5	5	5	5
$\tilde{n}_x = \frac{n_g}{\sum_{i=1}^{n} \frac{1}{n_{x_i}}}$	4	1.875	1.76	7.5	13.33

σ^2_g= 50; σ^2_{ga}= 200; σ^2_{gl}= 100; σ^2_{gal}= 10; σ^2_{plot}= 300;

$$\sigma^2_{p_n} = \sigma^2_g + \frac{\sigma^2_{ga}}{\tilde{n}_a} + \frac{\sigma^2_{gl}}{\tilde{n}_l} + \frac{\sigma^2_{gal}}{\tilde{n}_{al}} + \frac{\sigma^2_{plot}}{\tilde{n}_{alr}} = 50 + \frac{200}{4} + \frac{100}{1.875} + \frac{10}{7.5} + \frac{300}{13.33} = 177.17$$

$$H^2_{Holland} = \frac{50}{177.17} = 0.28$$

As seen here, the calculation is not as straightforward as in the previous methods presented but provides a more accurate estimate due to the use of harmonic means. It can be deduced that the 28% of phenotypic variation (49 out of 177.17 trait units) can be attributed to genetic differences.

"Piepho's" broad-sense heritability method

Given the regularity in which plant breeding programs face unbalanced data, alternative and more robust methods for estimating heritability been developed (Holland et al., 2003; Piepho and Möhring, 2007). The idea behind these methods is to obtain the non-genetic variance from the mean variance of a difference of two genotypic BLUEs, rather than attempting to deduce error variance from the plot error variance (divided by a factor that is a function of years, locations and replicates). The advantage of this method is it accounts well for unbalanced data since, as the squared standard errors or variance of the genetic estimates vary in size according to replication level, so the unbalanced data is directly account for. The disadvantage is that it requires at least two measurements of each individual to obtain a standard error, because the method is based on BLUEs (fixed effects):

$$H^2_{Piepho} = \frac{\sigma^2_g}{\sigma^2_g + (\bar{v}^{BLUE}_{\Delta}/2)} \sim \frac{\sigma^2_g}{\sigma^2_g + (\overline{EV}/2)}$$

Where: "σ2" refers to variance, "g" to genotype, $\bar{v}^{BLUE}_{\Delta}$ to mean variance of a difference of two genotypic BLUEs and $\overline{EV}$ is the average variance of genotype BLUEs.

Example: Assume MET run over four years across three locations in each year, for two reps in each year-location combination. After running a mixed model with genotypes fitted as fixed effects and the rest of terms as random (see data and script available in additional material) the following parameters are obtained:

genotype	mu (BLUE)	$(s.e.)^2$
1	80	50
2	56	60
3	240	75
4	100	80
5	180	70
$\frac{\sum_{i=1}^{n} \boldsymbol{s.e.}}{\boldsymbol{n}}$	-	67

σ^2_g= 50; σ^2_a= 200; σ^2_l= 100; σ^2_{al}= 10; σ^2_{plot}= 300 ; $\bar{v}_{\Delta}^{BLUE} = 67$

$$H^2_{Piepho} = \frac{50}{50 + (\frac{67}{2})} = 0.59$$

As seen here, the calculation is straightforward but it requires each individual to have more than one repetition in order to estimate a standard error for the estimates and requires two models to be fitted, one where the genetic term is fitted as fixed (to estimate $\bar{v}_{\Delta}^{BLUE}$), and another where the genetic term is fitted as random (to estimate σ^2_g). It can be deduced that the 59% of the phenotypic variation can be attributed to genetic differences.

Method 5. "Walsh and Lynch" broad-sense heritability method

This method proposed by Walsh and Lynch (2018) is also known as the BLUP-BLUE regression method and takes advantage from the fact that BLUP are shrunk by the factor $\sigma^2_e/\ \sigma^2_g$ whereas the BLUEs are not shrunk. This can provide an estimate of heritability a similar way to the Cullis method does [$1 - (\sigma^2_e/\ \sigma^2_g)$]. In this method, the regression coefficient between the BLUP and BLUE turns out to be the inverse of the shrinkage parameter [$\beta = 1 - (\sigma^2_e/\ \sigma^2_g) = H^2$]. This method is robust but one disadvantage is the need to fit the linear model twice (genotypes fitted as fixed and random effects), which can be computationally intensive under certain scenarios.

$$BLUP = \alpha + H^2\ BLUE$$

Where: "α" refers to the intercept, "H2" refers to the slope of the regression.

Example: Assume a multienvironment trial (MET) run in 4 years in 3 location in each year, and 2 reps in each year-location combination. After running a mixed model with genotypes fitted as random effects and the rest of terms as desired (see data and script available in additional material) we have come up with the following parameters:

genotype	Environment	u (BLUP)	b (BLUE)	mu	b-mu (BLUE_scaled)
1	**1**	0.501	5.107	4.47	0.628
2	**1**	-0.004	4.478	4.47	-0.0009
3	**1**	-0.784	3.499	4.47	-0.980
4	**1**	0.006	5.037	4.47	0.010
5	**1**	0.474	4.536	4.47	0.557
....	...	...	...	...	...

$$b = \frac{cov(BLUP,BLUE)}{var(BLUE)} = \frac{0.14}{0.17} = 0.82 = H^2$$

As seen here, the calculation is straightforward, but it requires genotypes to be modeled as both fixed and random. It can be deduced that the 82% of the phenotypic variation can be attributed to genetic differences.

Conclusion and recommendations

This manual has set out the correct interpretation of heritability alongside some common misconceptions to be avoided. Additionally, an overview was provided of the features, advantages and disadvantages of some of the more robust heritability calculation methods in order to promote the adoption of common and transparent methods among breeding programs. Among those, the Cullis method (Cullis et al., 2006) was recommended as a robust method to account for unbalanced datasets. In addition, the Piepho and Walsh & Lynch methods are also considered robust but require additional considerations.

Bibliography

Cullis, B.R., F.M. Thomson, J.A. Fisher, A.R. Gilmour, and R. Thompson. 1996. The analysis of the NSW wheat variety database. I. Modelling trial error variance. Theor. Appl. Genet. 92:21–27. doi:10.1007/BF00222947

Cullis, Brian R., Alison B. Smith, and Neil E. Coombes. "On the design of early generation variety trials with correlated data." Journal of agricultural, biological, and environmental statistics 11.4 (2006): 381.

Falconer, D.S., and T.F.C. Mackay. 2005. Introduction to quantitative genetics. 4th ed. Pearson Prentice Hall, Upper Saddle River, NJ.

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martínez. 2003. Estimating and interpreting heritability for plant breeding: An update. Plant Breed. Rev. 2003:9–112. doi:10.1002/9780470650202.ch2

Knight, Robert L. "Dictionary of genetics, including terms used in cytology, animal breeding and evolution." Dictionary of genetics, including terms used in cytology, animal breeding and evolution. (1948).

Kor Oldenbroek and Liesbeth van der Waaij, 2015. Textbook Animal Breeding and Genetics for BSc students. Centre for Genetic Resources The Netherlands and Animal Breeding and Genomics Centre, 2015. Groen Kennisnet: https://wiki.groenkennisnet.nl/display/TAB/

Lourenço, V. M., et al. "A robust DF-REML framework for variance components estimation in genetic studies." Bioinformatics 33.22 (2017): 3584-3594.

Oakey, H., A. Verbyla, W. Pitchford, B. Cullis, and H. Kuchel. 2006. Joint modeling of additive and non-additive genetic line effects in single field trials. Theor. Appl. Genet. 113:809–819. doi:10.1007/ s00122-006-0333-z

Piepho, H.-P., and J. Möhring. 2007. Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888. doi:10.1534/genetics.107.074229

Schmidt, P., et al. "Estimating Broad-Sense Heritability with Unbalanced Data from Agricultural Cultivar Trials." Crop Science 59.2 (2019): 525-536.

Schmidt, Paul, et al. "Heritability in plant breeding on a genotype-difference basis." Genetics 212.4 (2019): 991-1008.

Smith, A., B. Cullis, and A. Gilmour. 2001. The analysis of crop variety evaluation data in Australia. Aust. N. Z. J. Stat. 43:129–145. doi:10.1111/1467-842X.00163

Walsh, B., and M. Lynch. 2018. Evolution and selection of quantitative traits. 1st ed. Oxford Univ. Press, Oxford, UK. doi:10.1093/ oso/9780198830870.001.0001

Glossary

BLUP: Best linear unbiased predictor. Statistical estimate for a random effect with distribution $u \sim MVN(0, \Sigma\sigma^2_u)$ being Σ a relationship matrix among the levels of the random effect. Please see the EiB manual on BLUE vs BLUP to understand better the concepts of BLUE and BLUP.

BLUE: Best linear unbiased estimator. Statistical estimate for a fixed effect with distribution $\beta \sim MVN(\hat{\beta}, X'X^{-1})$ being Σ a relationship matrix among the levels of the random effect. Please see the EIB manual on BLUE vs BLUP to understand better the concepts of BLUE and BLUP.

REML: Restricted maximum likelihood. Statistical methodology for estimating variance components by maximizing the probability of the having variance components with certain values given the observed data for a response variable (i.e. a trait phenotype).

Harmonic mean: In mathematics, the harmonic mean (sometimes called the subcontrary mean) is one of several kinds of average, and in particular one of the Pythagorean means. Typically, it is appropriate for situations when the average of rates is desired. The harmonic mean can be expressed as the reciprocal of the arithmetic mean of the reciprocals of the given set of observations.

Random effect: Name assigned a covariate that aims to be fitted with the properties of a random variable with distribution $u \sim MVN(0, \Sigma\sigma^2_u)$ being Σ a relationship matrix among the levels of the random effect. Please see the "How to" tutorial on BLUE vs BLUP to understand better the concepts of BLUE and BLUP.