Check for updates

Received: 14 June 2020 | Revised: 4 February 2021 | Accepted: 8 February 2021

DOI: 10.1002/ece3.7389

ORIGINAL RESEARCH

Ecology and Evolution Open Access WILEY

Estimating heritability in honeybees: Comparison of three major methods based on empirical and simulated datasets

Hélène Jourdan-Pineau[1,2,3,4] | Gaëlle Antoine[1,4] | Julien Galataud[1,4] | Hélène Delatte[1,4] | Christophe Simiand[1,4] | Johanna Clémencet[1,4]

[1]CIRAD, UMR PVBMT, Saint-Pierre, France

[2]ASTRE, CIRAD, INRAE, Univ Montpellier, Montpellier, France

[3]CIRAD, UMR ASTRE, Montpellier, France

[4]UMR PVBMT, Université de La Réunion, St Denis, France

Correspondence
Hélène Jourdan-Pineau, CIRAD, UMR ASTRE, Montpellier, France.
Email: helene.jourdan@cirad.fr

Funding information
Centre de Coopération Internationale en Recherche Agronomique pour le Développement; Conseil Régional de La Réunion; European Agricultural Fund for Rural Development

Abstract

The genetic contribution to phenotypic variation (namely the heritability) affects the response to selection. In honeybee, the haplodiploid sex determination does not allow the straightforward use of classical quantitative genetics methods to estimate heritability and genetic correlation. Nevertheless, specific methods have been developed for about 40 years. In particular, sibling analyses are frequently used with three main methods: an historical model using the average colony relatedness, a half-sib/full-sib model, and the more recent animal model. We compared those three methods using experimental and simulated datasets to see which performs the best. Our experimental dataset is composed of 10 colonies with a total sample of 853 workers. All individuals were genotyped to reconstitute the pedigree, and phenotypic traits were measured: labial palpus and wing cubital veins lengths. We also simulated phenotypic datasets with varying levels of heritability, common environment effect, and genetic correlation between traits. The simulation approach showed that the average colony relatedness was highly biased in presence of common environment effect whereas the half-sib/full-sib and the animal model gave reliable estimates of heritability. The animal model provided the greatest precision in genetic correlations. Using this latter method, we found that wing vein lengths had high heritabilities whereas the palpus length had lower heritability due to larger environmental variance and/or measurement error. Finally, significant genetic correlations among measured traits indicate that they do not evolve independently.

KEYWORDS

Apis mellifera, genetic correlation, heritability, morphology, quantitative genetics, simulations

1 | INTRODUCTION

Phenotypic change in response to selection directly depends on the amount of phenotypic variation and whether this variation is transmitted from parents to offspring. Hence, predicting the evolutionary potential of a trait can be done by estimating the genetic component of the phenotypic variance V_P (namely the heritability) which is calculated from the phenotypic resemblance of related individuals in quantitative genetic analyses (Lynch & Walsh, 1998). More precisely, the genetic variance (V_G) can be partitioned into an additive component (V_A) as well as non-additive components: the dominance variance (V_D, due to the interaction of alleles at the same locus) and

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

the epistatic variance (due to interaction of alleles at different loci). The response to selection depends only on the additive part of the genetic variance as indicated by the (univariate) breeder's equation $R = h^2 S$, where R is the expected change in trait means per generation, h^2 the narrow-sense heritability and S the selection differential (Lush & Hubbs, 1945). Quantitative genetics analyses allow breeders to improve economically important characteristics in animals or plants and evolutionary biologists to understand or predict trait evolution.

Similarly, studies of heritability in the honeybee, *Apis mellifera*, had the primary goal of increasing honey production, disease resistance, and decreasing swarming and aggression (Koffler et al., 2017). Such economical concerns had motivated the development of specific solutions to apply quantitative genetics analyses to this haplodiploid polyandrous species. First, the morphological, physiological, and behavioral differences between castes preclude the use of straightforward parent–offspring regression approach (but coloration traits can be analyzed, e.g., Szabo & Lefkovitch, 1992). A noticeable exception is the Cape honeybee, *Apis mellifera capensis*, in which laying workers produce female workers, allowing parent–offspring regression to be performed (Brandes, 1988; Le Conte et al., 1994; Moritz & Hillesheim, 1985; Moritz & Klepsch, 1985). More importantly, the general framework of quantitative genetics is best developed for diploid organisms, reproducing randomly. If noncompliance with the last assumption has rather minor consequences on heritability estimates, the asymmetrical inheritance of parental genomes, due to haplodiploid sex determination, greatly complicates analyses.

For the last 40 years, adaptations of existing sibling analyses methods and protocols available for diploid species have been proposed. In 1977, Rinderer proposed to use artificial insemination of queens to control relatedness of the workers (Rinderer, 1977). In the case of single-drone inseminations, a colony consists of a full-sib family of workers with known relatedness $r = \frac{3}{4}$ whereas in the case of multiple-drone inseminations (with more than 20 drones), the colony approximates a diploid–diploid half-sib system (with $= \frac{1}{4}$) (e.g., Collins et al., ,1984, 1987; Harbo, 1992; Rinderer et al., 1983). Heritability is then estimated from the intracolony correlation t (also called the sibling correlation) such that $h^2 = \frac{t}{r}$.

Rinderer's method was soon improved to avoid time-consuming and expensive artificial insemination. In 1983, Oldroyd and Moran published a general formulae to calculate the ***average colony relatedness*** when a queen is mated with m drones, $r = \frac{1}{4} + \frac{1}{2m}$ (also published by Laidlaw & Page, 1984 or Milne & Friars, 1984). This approach has been widely used (until 2012, Goudie et al., 2012) even if it was early recognized that it produced upward bias in h^2 estimates (Diniz-Filho et al., 1993; Goudie et al., 2012; Melo et al., 1997; Oldroyd & Moran, 1983; Oldroyd et al., 1991; Poklukar & Kezić, 1994). Indeed, sib workers within a colony are raised in a common environment which further increases phenotypic resemblance beyond genetic effects (Oldroyd et al., 1991). Actually, this common environment is similar to a maternal (phenotypic) effect in which sibling's phenotypes are influenced by the maternal phenotype and by the environment provided by the mother.

When paternal lineages within colonies can be artificially manipulated (thanks to controlled fertilization) or identified based on genotyping, a typical ***half-sib/full-sib*** design can be applied. Liu and Smith derived heritability calculation of this sib design (and three other experimental designs) in the case of one male mated with several females (Liu & Smith, 2000). The reverse situation (one female-queen mated with several male-drones) is the natural situation in the honeybee and this leads to different formulae, as pointed out by Fjerdingstad (Fjerdingstad, 2005; Harpur et al., 2014; Laloi & Pham-Delegue, 2010).

With molecular genotyping, the additive genetic relationship matrix among individuals can be derived from the pedigree and included as random factor into the so-called ***animal model***, a linear mixed model (Wilson et al., 2010). This approach is more flexible than the ***half-sib/full-sib*** method since it considers all known degrees of relatedness between measured individuals and it does not require a balanced breeding design. However, the main issue when using the ***animal model*** is the inversion of the additive genetic matrix (which is not diagonal) in haplodiploid species. Bienefield et al. (Bienefeld et al., 2007 applied in Costa-Maia et al., 2011; Faquinello et al., 2011; Pernal et al., 2012) proposed a first solution using an approximation which was further improved by Brascamp and Bijma (Brascamp & Bijma, 2014; Brascamp et al., 2016). Even more recently, Bernstein et al. (Bernstein et al., 2018) published a new algorithm to facilitate the previous method on large datasets. Another approach to invert the additive genetic matrix is based on methods specifically developed for sex chromosome inheritance, which is functionally equivalent to haplodiploidy (Bohidar, 1964; Crow & Roberts, 1950; Crozier, 1970; Fernando & Grossman, 1990). In haplodiploid species, a mother contributes to the genome of both sons and daughters, whereas a father contributes only to the genome of daughters. Similarly, in sex chromosome, genes linked to the X-chromosome are transmitted to daughters and sons, whereas genes linked to the Y-chromosomes are only transmitted to sons. This method of matrix inversion is implemented in the R package **nadiv** (Wolak, 2012), which has been applied to wasps but never to honeybees (Sheehan et al., 2017).

However, many traits under selection are genetically correlated and will not evolve independently from each other, due to linkage or pleiotropy (one loci influencing several traits). In such case, estimating genetic correlation between traits allows a better understanding of trait evolution. They are calculated following the same approaches as heritability estimates but require larger sample size to obtain reliable estimates (around 1,000 individuals when considering only two traits, Brown, 1969).

These three sibling analysis methods, ***average colony relatedness, half-sib/full-sib, and animal model***, have allowed estimating heritability and genetic correlations of a vast number of traits in honeybee (e.g., morphology, behavior, life-history traits). Koffler et al. (2017) reviewed published estimates in honeybee (and other bee, ant, and wasp species), and they showed that heritability estimates varied between trait types. This is expected from Fisher's fundamental theorem saying that traits more closely related to fitness should

be more submitted to selection which decreases genetic variation (Fisher, 1958). However, their meta-analysis did not detect any effect of the analytical methods used (among which sibling analyses and parent–offspring regression) on heritability estimates. This result, contrasting with previously published studies (e.g., Postma, 2014), could either be due to restricted statistical power or to biases inherent to meta-analysis: datasets diverge on several factors (analytical method, sample size, and traits) resulting in confounded effects. In addition, a few studies compared different methods in the same honeybee dataset and found that those methods yielded heritability estimates in the same range (Brandes, 1988; Moritz, 1985; Moritz & Klepsch, 1985). Genetic correlation estimates are scarce (14 studies only in honeybee), and half of the estimates are not significantly different from 0 or were not provided with any significance tests or standard errors (Koffler et al., 2017). Overall, previous works do not allow to draw any firm conclusion on the performance of different methods, sample size, or experimental design to estimate genetic correlation and heritabilities. One option is to use simulation studies to compare available methods (as well as sample size and design) based on their accuracy of estimates and on their statistical power to detect heritability and genetic correlation (De Villemereuil et al., 2013; Holand & Steinsland, 2016; Morrissey et al., 2007).

In this paper, we used both empirical and simulated datasets to compare the performance of three sibling analysis methods, ***average colony relatedness, half-sib/full-sib, and animal model,*** in the estimation of heritability and genetic correlations. Our empirical dataset is composed of three morphological traits measured on 853 workers from ten colonies of *Apis mellifera unicolor* sampled in two islands of the South-West Indian Ocean. The three traits (associated with mouthpart and forewing morphometry) are used to discriminate between species or populations, indicating relatively fast evolutionary rates (Cornuet et al., 1975; Ruttner, 1988; Ruttner et al., 1978). We also generated simulated phenotypic datasets with varying levels of heritability, genetic correlation, and common environment effect.

2 | MATERIAL AND METHODS

2.1 | Empirical dataset

Honeybee workers were sampled in Mauritius and La Reunion, two islands of the Mascareignes archipelago situated in the South-West Indian Ocean. In this area, the local honeybee subspecies is *A. m. unicolor* belonging to the African lineage with recent import of the *A. m. carnica*, *ligustica*, *and mellifera* subspecies belonging to the European lineage (Techer, Clémencet, Simiand, Turpin, et al., 2017). La Reunion and Mauritius honeybee populations are genetically distinct (Techer, Clémencet, Simiand, Turpin, et al., 2017).

We sampled 95 worker honey bees per colony, from the frames of the hives, in a total of ten colonies. Those workers are a mixture of full-sisters and half-sisters (sharing the same queen mother but different fathers). Five colonies were sampled on Mauritius in October 2014, whereas five colonies were sampled on La Reunion from 2011 to 2018 (Appendix S1). Individuals were kept in 95% ethanol at −80°C until they were processed.

Patrilines were determined by genotyping autosomal microsatellite markers on DNA extracted from the femur (as described in Techer, Clémencet, Simiand, Preeaduth, et al., 2017). Eight microsatellite markers (A24, AC306, AP55, A289, A8, AP33, and AP66) were selected from a larger list (17 loci) to be as variable and informative as possible (Solignac et al., 2003; Techer, Clémencet, Simiand, Turpin, et al., 2017). PCR reactions were multiplexed and performed in 10-µl volumes containing 5 µl of Master Mix Type-it 2× Qiagen, 0.2 µl of each fluorescent-labeled primer at 20 pmol/µl and 1 µl of DNA at 5 ng/µl. All programs started with a denaturing cycle at 94°C for 15 min, followed by 35 cycles of 30 s at 94°C, 30 s at 52°C, 45 s at 72°C, and a final elongation at 72°C for 20 min. The samples were run through a DNA sequencer ABI Prism 3130XL, and alleles were scored using GeneMapper 4.0 (Applied Biosystems). Only fully genotyped workers have been considered in the reconstruction analyses. In addition, workers presenting incoherent genotype (potential drifting workers from other hives) have been excluded from the dataset. Thus, 424 individuals were kept from Mauritius (exclusion of 49 individuals) and 429 individuals from La Reunion (exclusion of 46 individuals) (Appendix S2). Patriline reconstruction was carried out on the basis of allelic frequencies previously estimated for those insular populations (Techer, Clémencet, Simiand, Turpin, et al., 2017) using MATESOFT software (Moilanen et al., 2004). Patriline reconstruction is facilitated by the haplodiploid determination of sex (Estoup et al., 1994). We found a mean number of 38 patrilines per colony and confirmed that there was only one queen per colony.

2.2 | Morphometric measurements

For each worker, the mouthparts and the right forewing were dissected and digitally photographed. The right forewing of each bee was cut at its base and mounted in water between micrometer blade and cover. The mouthparts were dissected and then mounted on a micrometer slide with a strip of adhesive tape.

Pictures were acquired using a video camera, and measurements were made on the AxioVision SE4 software 4.7. To study the variation in size of the mouthparts, we measured the length of the long segment of the right palpus (Appendix S3) which is not influenced by the extension of the proboscis and hence more repeatable than the proboscis (Morimoto, 1968). For the wings, we measured the length of the cubital veins A and B and computed the cubital index (CuI) as the ratio $\frac{A}{B}$, an index informing on the shape of the wing. This index allows to discriminate between species, subspecies, and even populations (Cornuet et al., 1975; Ruttner, 1988; Ruttner et al., 1978). The palpus length and cubital cell of 36 individuals were measured 4 times with two montages and two measurements for each of these montages. Measurement errors were mainly due to slide mounting, and it was almost null for the cubital cell (A, B, and CuI) and less than 2% for the palpus length. Therefore, experimental noise would not artificially increase the residual variance and bias heritability.

All measured traits (except for CuI) significantly differed between colonies and between islands (Appendix S4). Morphological measurements were globally larger on La Reunion than on Mauritius. In addition, a negative phenotypic correlation between the A and B veins ($r = -0.14$, CI: [−0.21, −0.08]) and a positive correlation between A vein and palpus length ($r = 0.32$, CI: [0.26, 0.38]) were observed (Appendix S5).

2.3 | Quantitative genetics analysis methods

Three commonly used methods based on sibling analysis were compared for their performances in estimating heritabilities and genetic correlations: (a) a simple linear model considering a colony effect and an ***average colony relatedness*** between individuals (Oldroyd & Moran, 1983); (b) a nested model of patrilines in the colonies equivalent to a ***half-sib/full-sib*** approach (Fjerdingstad, 2005); and (c) an ***animal model*** based on the pedigree of individuals (Sheehan et al., 2017).

In the three methods, dominance variance (V_D) cannot be separated from additive variance (V_A) due to haplodiploidy (Fjerdingstad, 2005; Liu & Smith, 2000). Thus, only the broad sense heritability is estimated: $H^2 = \frac{V_G}{V_P}$. In addition, those methods do not model the epistatic variance which will contribute to the genetic variance. This epistatic variance is very difficult to estimate and assumed to be negligible, at least for diploid organisms (Lynch & Walsh, 1998).

The first method does not require the identification of patrilines of the measured workers but an estimate of the number of efficient matings (m) is needed. Heritability is then calculated as follows:

$H^2 = \frac{t}{r}$, where r is the average colony relatedness $r = \frac{1}{4} + \frac{1}{2m}$ (Oldroyd & Moran, 1983)

and t the intracolony correlation $t = \frac{V_{col}}{V_{col} + V_R}$ with V_{col} the variance among colonies and V_R the (residual) variance within colonies (Lynch & Walsh, 1998). Note that the estimate of r changes little when m is greater than 8, which appears to be usually the case for freely mating queens (Oldroyd & Moran, 1983).

The genetic correlation (r_G) is given by:

$r_G = \frac{\mathrm{cov}_{\mathrm{col}\, t_1 t_2}}{\sqrt{V_{\mathrm{col}t_1} * V_{\mathrm{col}t_2}}}$ with $\mathrm{cov}_{\mathrm{col}\, t1t2}$ the covariance between traits t_1 and t_2 among colonies.

The second method is based on the identification of the patriline of each worker and applies a mixed-effects model with patrilines nested within colonies as random effect (Fjerdingstad, 2005). Heritability is calculated according to the following formula:

$H^2 = \frac{2 \times V_{pat}}{V_{pat} + V_{CE} + V_R}$ (with V_{pat} the variance among patrilines, V_{CE} the variance associated with the common environment of the colony and V_R the residual variance) owing that $V_{pat} = \frac{1}{2}V_A + \frac{1}{2}V_D$. Note that the coefficient for the dominance variance is $\frac{1}{2}$ because all paternal half-sibs get the same alleles from their father.

The genetic correlation is estimated by: $r_G = \frac{\mathrm{cov}_{\mathrm{pat}\, t_1 t_2}}{\sqrt{V_{\mathrm{pat}\, t_1} * V_{\mathrm{pat}\, t_2}}}$

with $\mathrm{cov}_{\mathrm{pat}\, t1t2}$ the covariance between traits t_1 and t_2 among patrilines.

The third method is based on a mixed-effect linear model where the pedigree is used to derive the additive genetic relationship matrix among individuals, fitted as random effect. To create the inverse of the additive genetic matrix required by **asreml-R**, we used the **makeS** function of the **nadiv** package implemented in **R** (Wolak, 2012). This function returns the inverse of the additive genetic relationship matrix for the sex chromosomes since haplodiploidy is equivalent to total sex-linkage of all genes. Here diploid females are coded as the heterogametic sex whereas haploid males are the homogametic sex. A "colony" (common environment) random effect is added to distinguish the effect of common environment from the genetic effect of the queen. Thus, heritability can be calculated as follows:

$H^2 = \frac{V_A}{V_A + V_{CE} + V_R}$ where V_A is the variance associated with the additive genetic matrix (remember that in our case, V_A is confounded with V_D), V_{CE} the variance associated with the common environment of the colony and V_R is the residual variance.

The genetic correlation between two traits (t_1 and t_2) is calculated based on bivariate animal models according to the formula:

$r_G = \frac{\mathrm{cov}_{G_{t_1 t_2}}}{\sqrt{V_{A_{t_1}} * V_{A_{t_2}}}}$ with cov_G the genetic covariance between traits t_1 and t_2.

In the second and the third methods, the amount of variance attributable to the common environment effect (CE^2) can be calculated as follows: $CE^2 = \frac{V_{CE}}{V_P}$.

To test whether heritability and genetic correlation estimates are significantly different from 0, log-likelihood ratio tests between nested models (with and without colony/patriline/additive matrix/genetic covariance) were performed.

2.4 | Quantitative genetics analyses of the empirical dataset

We applied the three methods described above to calculate heritabilities and genetic correlations between our measured traits (A,

B, CuI, and palpus). The **ASReml-R** package (Gilmour et al., 2009) was used to perform variance decompositions for all three methods even if **ASReml-R** was only required for the ***animal model*** whereas the other two methods could be performed with **lme4** or **nlme** R packages. Estimates were calculated using the whole dataset (853 individuals). Year and island were added as fixed effects in the statistical models. We assessed the significance of the fixed effects with a conditional Wald *F* test.

Using the ***animal model***, we also tested if the additive, common environment, and residual variances differed between islands using a bivariate model. Log-likelihood ratio tests allowed testing if variances were significantly different between islands. The approximate standard errors of heritabilities and genetic correlations were estimated with the **nadiv** R package, applying the delta method (Lynch & Walsh, 1998; Wolak, 2012).

2.5 | Quantitative genetics analyses of simulated datasets

We performed simulations to test the performances of the three methods in estimating heritability and genetic correlations. Our hypotheses were that (a) the ***average colony relatedness*** method would provide biased estimates in presence of common environment effect since this colony effect would be confounded with the genetic effect (Kruuk & Hadfield, 2007); (b) the ***animal model*** would allow to deal with unbalanced design and with various degree of relatedness between pairs of individuals. To test the first hypothesis, we simulated phenotypic datasets including a common environment effect resulting from the fact that workers from one colony share the same maternal environment. To test the second hypothesis, we used three different pedigrees: (a) our experimental pedigree which is an unbalanced pedigree, (b) a balanced pedigree with similar sample size as our experimental pedigree (840 workers in 10 colonies), and (c) a pedigree with added relatedness between queens (a modified version of our experimental pedigree).

2.5.1 | Heritability

Following the approach of Morrissey et al. (Morrissey et al., 2007), we first simulated a dataset of individual phenotypic values such that $Y_i = \mu + A_i + CE_k + \varepsilon_i$ with μ, the average phenotype in the population (arbitrarily set to 0); A_i, the genetic value of the individual *i* (normally distributed assuming an additive genetic variance V_A); CE_k, the common environment effect (normally distributed with V_{CE} variance); and ε_i, the residual variation (normally distributed with V_R variance). The genetic values A_i were generated according to the pedigree (see details below) and V_A using the **grfx** function of the **nadiv** package (Wolak, 2012). We did not simulate any epistatic variance.

In each scenario studied, the heritability (h^2) and common environment (CE^2) values varied according to four fixed values: 0 (absence), 0.1, 0.3, and 0.5 thus making it possible to compare 16 different phenotypic scenarios.

For each scenario, 1,000 phenotypic datasets were simulated using one of the three tested pedigrees. The first one is our unbalanced experimental pedigree (with 853 workers in 10 colonies). The second one is a simulated balanced pedigree. This simulated pedigree was built to be of similar size and design to our experimental pedigree: It is composed of 840 phenotyped individuals from 10 queens each crossed with 42 males (against an average of 38 males in our real pedigree), with a full-sib family size of two workers. The third pedigree was a modified pedigree from our experimental dataset (i.e., with 853 individuals), where three queens of three different colonies (out of the 10 colonies in the dataset) were considered to be full-sisters, therefore adding relatedness relationship between those three colonies.

Therefore we generated 48,000 datasets (16 phenotypic scenarios × 3 pedigrees × 1,000 datasets). We estimated the heritability, confidence intervals, and significance of the "genetic" effect (corresponding to the additive genetic matrix) of each simulated phenotypic dataset with the three previously described methods (***average colony relatedness***, ***half-sib/full-sib,*** and ***animal model***), except for datasets generated with the third pedigree (with related queens) for which we compared only the half-sib/full-sib and animal model.

For each dataset, we computed the mean and 2.5%, 25%, 75%, and 97.5% quantiles to quantify the bias and the dispersion of heritability estimates. We also calculated the power to detect heritability as the percentage of simulated datasets with a significant genetic effect. However, when estimates are biased, power is not a suitable parameter to assess the performance of a method and should be replaced by the coverage of 95% of confidence intervals (Bolker, 2008). Confidence intervals were calculated assuming a normal distribution of estimates ($h^2 \pm 1.96 \times SE$). Coverages were computed as the proportion of datasets in which simulated value of heritability was within the 95% confidence interval.

2.5.2 | Genetic correlation

Using the **grfx** function of the **nadiv** R package, we simulated two correlated traits with heritabilities being a combination of 0.1, 0.3, and 0.5 (hence nine combinations). We skipped the null value since there could not be genetic correlation between traits that are not heritable. We did not include any common environmental effect. We simulated only positive genetic correlations between traits, considering that the results would be symmetrical with negative genetic correlations. The simulated genetic correlation values were 0, 0.1, 0.3, and 0.5 (according to the genetic correlation estimates published in honeybee (Koffler et al., 2017)). The statistical performances of the three methods were evaluated as above (precision, power, and coverage).

TABLE 1 Estimated genetic parameters of the measured traits estimated (the cubital veins A and B, the cubital index (CuI) and the length of the long segment of the right palpus) using three methods: an ***average colony relatedness***, a ***half-sib/full-sib model*** (HS/FS), and an ***animal model***

Method	Trait	V_{col}		V_R	h^2	*SE* (H^2)	95% CI	
Colony relatedness	A	0.0002		0.0011	0.6628	0.3101	0.0550	1.2707
	B	0.0001		0.0005	0.8598	0.3717	0.1313	1.5883
	CI	0.0230		0.1082	0.6664	0.3118	0.0554	1.2775
	palpus	0.0054		0.0020	2.7669	0.4056	1.9719	3.5618
HS/FS		V_{pat}	V_{CE}	V_R	h^2	*SE* (H^2)	95% CI	
	A	0.0003	0.0003	0.0008	0.4315	0.0894	0.2562	0.6068
	B	0.0002	0.0001	0.0003	0.5821	0.0982	0.3897	0.7746
	CI	0.0305	0.0212	0.0763	0.4772	0.0883	0.3042	0.6502
	palpus	0.0009	0.0055	0.0011	0.2281	0.0949	0.0422	0.4141
Animal model		V_G	V_{CE}	V_R	h^2	*SE* (H^2)	95% CI	
	A	0.0006	0.0001	0.0007	0.4315	0.0894	0.2562	0.6068
	B	0.0003	0.0000	0.0002	0.5821	0.0982	0.3897	0.7746
	CI	0.0611	0.0059	0.0610	0.4772	0.0887	0.3034	0.6510
	palpus	0.0017	0.0051	0.0007	0.2282	0.0949	0.0422	0.4142

Note: Depending on the model, values for variance associated with colony (V_{Col}), variance associated with patriline (V_{Pat}, for the half-sib/full-sib model), genetic variance (V_G), common environment variance (V_{CE}), residual variance (V_R), heritability (H^2), and its standard error (*SE*) are presented. Year and Island were included as fixed effects. All genetic effects were significantly different from zero as tested by likelihood ratio tests.

TABLE 2 Genetic correlation, standard error (*SE*), and 95% confidence interval (CI) between pairs of measured traits (the cubital veins A and B, the cubital index (CuI) and the length of the long segment of the right palpus) estimated using three methods: an ***average colony relatedness***, a ***half-sib/full-sib*** model and an ***animal model*** (without island as random effect)

	Average colony relatedness				Half-sib/full-sib				Animal model			
Genetic correlation	Estimate	*SE*	95% CI		Estimate	*SE*	95% CI		Estimate	*SE*	95% CI	
A-B	0.17	0.34	−0.49	0.84	−0.57	0.07	−0.72	−0.42	−0.49	0.06	−0.61	−0.38
A-palpus	0.72	0.17	0.39	1.05	−0.46	0.07	−0.61	−0.32	0.21	0.07	0.07	0.36
B-palpus	0.05	0.34	−0.62	0.72	−0.49	0.07	−0.63	−0.36	0.02	0.07	−0.11	0.16
IC-palpus	0.42	0.29	−0.14	0.98	−0.07	0.10	−0.27	0.13	0.06	0.07	−0.09	0.21

3 | RESULTS

3.1 | Empirical dataset

All the studied traits showed significant heritability but estimates varied according to the method used (Table 1). The ***average colony relatedness*** method provided the highest heritability estimates whereas the ***half-sib/full-sib*** and the ***animal model*** gave strictly similar estimates from 0.25 for palpus length to 0.61 for the B length. Using the ***animal model***, we estimated the effect of common environment CE^2: low for the B vein and the cubital index (0.01 and 0.02, respectively), intermediate for A vein (0.15), and high for palpus length (0.65).

The three methods gave very different genetic correlation estimates (Table 2). According to the ***animal model***, the genetic correlations were significantly different from 0 between A and B veins (r_G = −0.49, 95% CI: [−0.61, −0.38], $p = 1.35 * 10^{-12}$) and also between A vein and the palpus length (r_G = 0.21, 95% CI [0.07, 0.36], p =.0051).

Using the ***animal model***, variance estimates were not significantly different between years ($p = 1$). Variance estimates differed between islands for palpus length but not for wing-related traits (Table 3). Indeed, palpus length heritability was lower on Mauritius (H^2 = 0.18) than on La Reunion (H^2 = 0.46). This difference was mostly due to an almost 10 times higher common environment variance on Mauritius.

3.2 | Simulations

Analyses carried out with the three methods with simulated datasets based on our experimental pedigree provided estimates close to the expected values in the absence of common environment effect (Figure 1, upper part). In the presence of a common environment effect (CE^2), the ***average colony relatedness*** largely overestimated heritabilities, resulting in h^2 values greater than 1, when the ***half-sib/full-sib*** and the ***animal model*** methods provided similarly reliable estimates (though with slight upward bias of around 8% for CE^2 = 0.5,

TABLE 3 Estimated genetic parameters of the measured traits (the cubital veins A and B, the cubital index (CuI) and the length of the long segment of the right palpus) estimated using a bivariate ***animal model***, with distinct parameters for each island

	Mauritius					Reunion					
Trait	V_G	V_{EC}	V_R	h^2	*SE* (H^2)	V_G	V_{EC}	V_R	h^2	*SE* (H^2)	*p*-value
A	0.0005	0.0005	0.0007	0.3104	0.1227	0.0007	0.0000	0.0007	0.4952	0.1335	0.1851
B	0.0003	0.0000	0.0002	0.5603	0.1299	0.0003	0.0000	0.0002	0.6196	0.0835	0.3539
CI	0.0619	0.0047	0.0654	0.4689	0.1151	0.0588	0.0037	0.0576	0.4899	0.1293	0.2748
palpus	0.0019	0.0080	0.0005	0.1859	0.1034	0.0015	0.0012	0.0009	0.4092	0.1733	0.0075

Note: We presented values for genetic variance (V_G), variance associated with common environment (V_{CE}) and residual variance (V_R), heritability (H^2), and its standard error (*SE*) for each island. *p*-values indicated if variances significantly differed between islands, as tested by likelihood ratio tests.

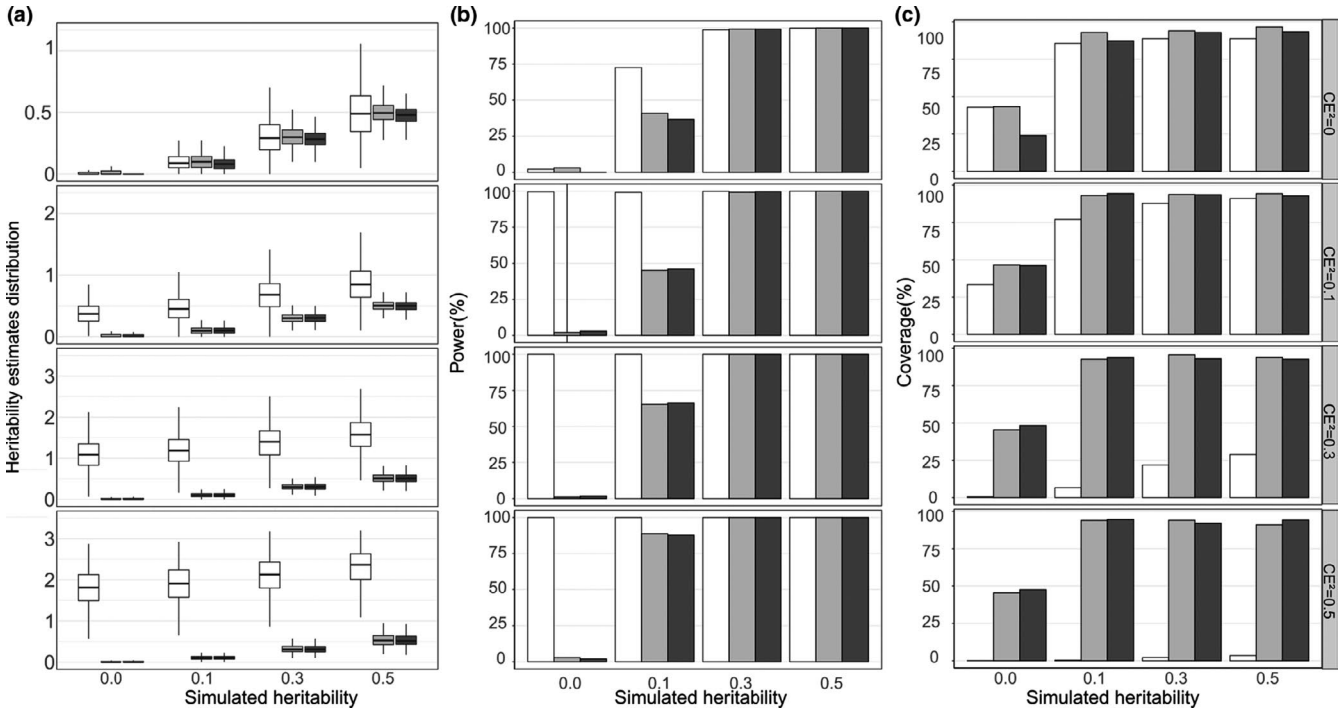


FIGURE 1 Performance of heritability estimates for the different models used (***average colony relatedness***, ***half-sib/full-sib***, and ***animal model***, white, gray, and black bars, respectively) depending on the levels of simulated heritabilities and of common environment effects (CE^2, vertical panels). (a) Distribution of heritability estimates. Boxes represent 25 and 75% quantiles, and the middle line represents the mean. Whiskers show 2.5% (bottom) and 97.5% (top) quantiles. Note that the scales of the Y-axis vary for the different levels of CE^2 to facilitate comparison between the three models. (b) Power to determine genetic effect and (c) coverage of 95% confidence interval of heritability estimates. Simulated datasets were based on our experimental pedigree

Figure 1). Similar results were found using a balanced pedigree (Appendix S6).

For the three methods (and the three tested pedigrees), the rate of false positive (probability to obtain a significant genetic effect whereas $h^2 = 0$) was low (<3%), except for the ***average colony relatedness*** method in presence of a common environment effect due to large bias in heritability estimates (as mentioned above, Figure 1. right side). Accordingly, coverage was very low for the ***average colony relatedness*** when the common environment effect was moderate to high ($CE^2 > 0.3$). The two other methods led to greater coverage but still below 95%, whatever the value of the true heritability. Finally, the statistical power was low (between 35% and 90%) for heritabilities of 0.1 but very high (>98%) for moderate to high heritabilities ($h^2 \geq 0.3$).

Using the simulated pedigree with related queens, we found that the ***half-sib/full-sib*** and the ***animal model*** methods yielded similar results as mentioned above for our empirical pedigree. We noticed only a very slight decrease in the precision of estimates for the ***half-sib/full-sib*** method (Appendix S7).

Concerning genetic correlations, the ***average colony relatedness*** method gave estimates with upward bias, low precision, and low power (Figure 2a). On the contrary, the ***half-sib/full-sib*** and the ***animal model*** methods gave fairly good estimates, in particular when the two considered traits had high heritabilities. However, the ***half-sib/full-sib*** method has a slight upward bias hence lower coverage of 95% CI (Figure 2a,c). The statistical powers of ***half-sib/full-sib*** and the ***animal model*** methods were low when genetic correlation was lower than 0.5 except when the two traits were highly heritable (Figure 2b). Note that the ***half-sib/***

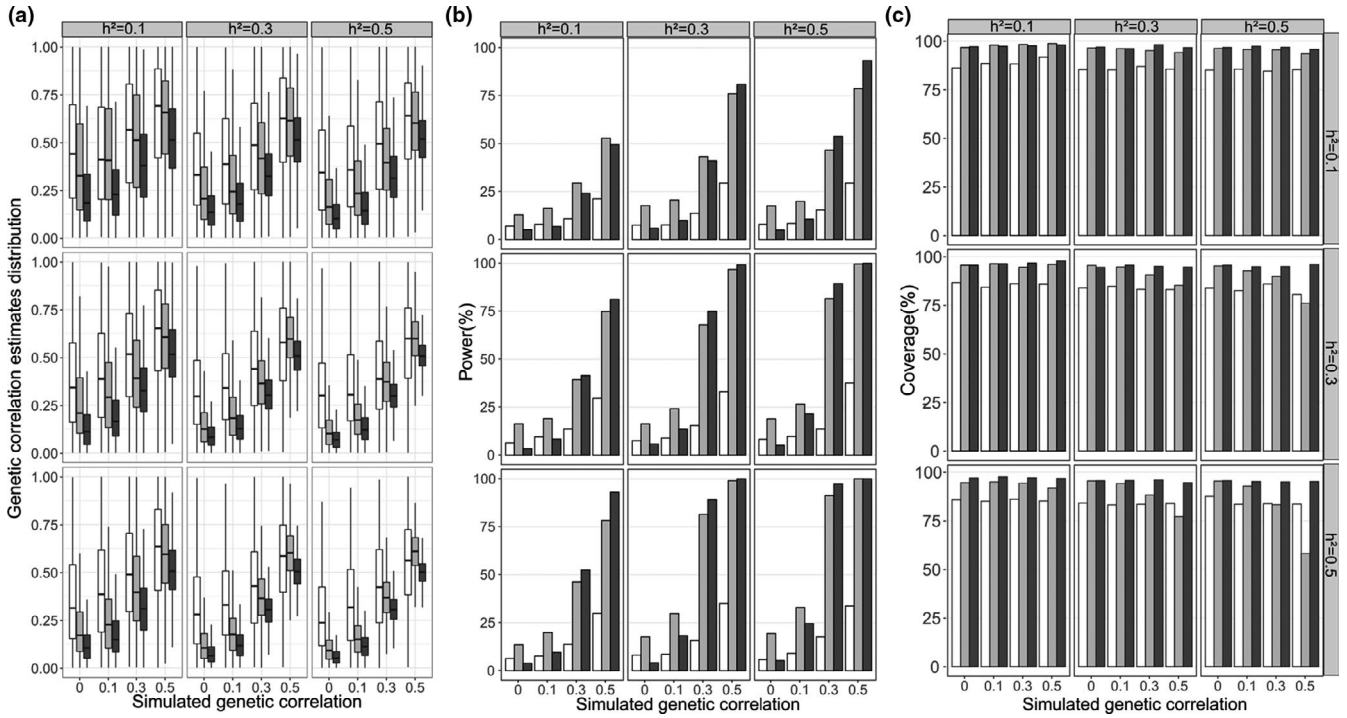


FIGURE 2 Performance of genetic correlation estimates for the different models used (***average colony relatedness***, ***half-sib/full-sib,*** and ***animal model***, white, gray, and black bars, respectively) depending on the levels of simulated genetic correlations and simulated heritabilities of the two traits (h^2 of trait 1 and trait 2, horizontal and vertical panels). (a) Distribution of genetic correlation estimates. Boxes represent 25 and 75% quantiles, and the middle line represents the mean. Whiskers show 2.5% (bottom) and 97.5% (top) quantiles. (b) Power to determine genetic correlation and (c) coverage of 95% confidence interval of genetic correlation estimates

full-sib method was subject to convergence failure (when using default convergence parameters Appendix S8). This was largely determined by the estimated heritabilities of the two correlated traits: failure to converge was high when heritabilities were low.

4 | DISCUSSION

We compared the performances of three sibling analysis methods, ***average colony relatedness, half-sib/full-sib, and animal model,*** in the estimation of heritability and genetic correlations. Of the three methods tested, the ***average colony relatedness*** relies the least on patriline reconstruction, since it is based on a mean number of patrilines per colony and the calculation of average relatedness is little impacted by the effective number of matings (Oldroyd & Moran, 1983). Hence, it is not necessary to genotype all individuals which constitutes a real practical and financial advantage. The two other methods were influenced by the quality of pedigree reconstruction which depends of genotyping errors, the presence of null alleles, or informativeness of genetic markers (Csilléry et al., 2006; Pemberton, 2008; Wang, 2006). Here, the number of patrilines estimated for each colony in our experimental dataset fluctuates according to the number microsatellite markers used: expectedly, with seven markers instead of eight, the number of patrilines estimates per colony was lower. Similarly, heritability estimates were lower with seven microsatellite markers (Appendix S9). Indeed, when the number of patrilines estimated is less than the actual number of fathers, half-sisters will be considered as full-sisters while their resemblance will be weaker than expected, thus leading to a lower heritability (Firth et al., 2015). However, the magnitude of underestimation associated with a lower number of microsatellite markers was small (typically between 3% and 15% depending on the considered trait). Those values are in line with already published biases associated with pedigree errors (Bérénos et al., 2014; Charmantier & Réale, 2005; Firth et al., 2015) and have little impact on general interpretations.

Despite its practical simplicity, the ***average colony relatedness*** method suffers from a major flaw: an upward bias in heritability estimates (even yielding estimates higher than 1) caused by common environmental rearing condition of the workers (Poklukar & Kezić, 1994). This major limitation had already been acknowledged before, and this bias was drastically reduced by cross-fostering of offspring workers into different rearing environment (Oldroyd et al., 1991). Another solution is to reduce the differences in rearing/developing conditions between colonies, using a "common garden" experiment (see, e.g. Moritz & Hillesheim, 1985; Oldroyd & Moran, 1983). Under those restricted conditions (no or controlled environmental effect), the ***average colony relatedness*** could constitute a convenient solution (although with less precise estimates).

Both the ***half-sib/full-sib*** and the ***animal model*** provided precise and unbiased heritability estimates with sufficient power (in particular with intermediate to high heritabilities), either on our

experimental pedigree or on the simulated pedigree. The similar performance of the two methods with unbalanced pedigree was contrary to our expectations. First, this may be due to the fact that our experimental design is not heavily unbalanced. Second, we used restricted maximum likelihood procedures that are relatively insensitive to unbalanced designs (contrary to ANOVA approaches). In such situation, using either of both methods are equally pertinent and people who are not familiar with the ***animal model*** (and the inversion of the additive matrix with appropriate tools) may prefer to use the ***half-sib/full-sib*** method.

In our experimental pedigree, we reconstructed a fairly simple pedigree (similar to a half-sib/full-sib design) and considered that fathers and queens were unrelated to each other. This assumption was reasonable since mating occurs during nuptial flight in congregation area where thousands of drones gather, from a large number of colonies in the surroundings, resulting in an almost perfect panmixis (Baudry et al., 1998). In such a situation, the copulation of the queen with one drone related to herself or with two drones related to each other is very unlikely. In addition, when colonies are sampled in distant areas, queens are less susceptible to be related.

However, colonies sampled in the same apiary (as is the case for two colonies of our dataset sampled in Le Baril) may be related, if they originated from the division of one hive into daughter colonies. To explore this situation, we simulated a pedigree with related queens. We expected the ***half-sib/full-sib*** method, which ignores such additional relationship, to be more biased than the ***animal model***, which takes into account all types of relationships between individuals. Surprisingly, this was not the case and both methods provided reliable estimates. This may be due to the fact that the added relatedness links were negligible, and it echoes the demonstration of Liu and Smith (2000) showing that moderate inbreeding may not notably bias the genetic parameters estimated by sib analysis. Therefore, in most cases, ***half-sib/full-sib*** and ***animal model*** methods may be used but the ***animal model*** will yield slightly more precise estimates (i.e., with the smallest RMSE). In addition, the ***animal model*** method can handle complex models which may be required to study colony traits (Bienefeld et al., 2007; Bienefeld & Pirchner, 1990) but also to deal with dominance (Wolak & Keller, 2014).

In the case of a haplodiploid species, the genetic variance estimated by the ***animal model*** is composed of an additive and a dominance component. According to Fisher's assumption, in population with large effective size, the variance of dominance is supposed to be negligible (Wolak & Keller, 2014). Indeed, the dominance effect is linked with the genetic background in which it is expressed (Fisher, 1958). Thus, in a population of infinite size, every possible genetic backgrounds are represented and the dominance effects average out to zero. In the case of La Reunion and Mauritius, the effective honeybee population size seems to be large (Techer, Clémencet, Simiand, Turpin, et al., 2017). Hence, the dominance variation, which are likely to contribute to our heritability estimates, will have a limited effect on the genetic variance of the population. One meta-analysis showed that the dominance variance is only important for domesticated species and is generally low (0.15, Wolak & Keller, 2014). More recently, Class and Brommer (2020) confirmed that the dominance variance represents a modest amount of genetic variance in a wild population of blue tits.

As mentioned above, it is theoretically possible to determine the dominance variance in an ***animal model*** using the dominance matrix which can be obtained using the **makeSd** function of the **nadiv R** package (designed for sex chromosome inheritance and thus appropriate for haplodiploid organisms) (Wolak, 2012). In a design like ours (half-sib/full-sib), the additive matrix is identical to the dominance matrix, which prevents to separate the two variances V_A and V_D. A more complex pedigree (with more complex kinship relationships) would resolve this constraint (Wolak & Keller, 2014).

Considering genetic correlation, our simulation approach clearly demonstrates that the ***average colony relatedness*** method is not suitable for such task even in the absence of common environment effect: estimates were biased and dispersed (except when the two traits were highly heritable), and the statistical power was very low. The ***half-sib/full-sib*** and the ***animal model*** were more appropriate with the latter less prone to false positive and providing more accurate estimates. We did not test the impact of sample size on the performances of the three methods (either on genetic correlation or heritability) but it has been regularly demonstrated that quantitative genetics require large sample size (at minimum 250 individuals and preferably over 500), in particular when estimating genetic correlation (which requires rather almost 1,000 individuals (Brown, 1969; Lynch & Walsh, 1998; De Villemereuil et al., 2013). In honeybee, most colonies are easily accessible and colony sizes are very large. This allows sampling a large number of individuals, families, and colonies, and hence, the only limitation to obtain robust estimates is the technical and financial ability to phenotype and genotype all sampled individuals.

According to the simulation results, we decided to discuss only the estimates provided by the ***animal model*** in our experimental dataset. This model yielded high heritability estimates for the four measured traits, which is consistent with the honeybee literature (Moritz & Klepsch, 1985; Mostajeran et al., 2006; Oldroyd et al., 1991; Poklukar & Kezić, 1994; reviewed in Koffler et al. 2017). In general, morphological traits are known to display higher heritabilities than fitness-related traits (Mousseau & Roff, 1987). The lower heritability of the palpus length could be in part due to errors of measurement contributing to residual variance. We also found that variance estimates are not significantly different for wing traits between the two islands indicating little influence of environmental factors. In contrast, palpus length is less heritable on Mauritius than on La Reunion due an almost 10 times larger common environment variance. Accordingly, the proboscis is the morphological character showing the largest geographic variability (Ruttner, 1988; Ruttner et al., 1978) supporting a great phenotypic plasticity. In addition, colonies came from more diverse environments on Mauritius compared to La Reunion where two areas were sampled twice (Le Barril and Ligne Paradis/Bassin Plat).

Our results indicated that the A vein was genetically correlated with the B vein and with the palpus. However, genetic correlation estimates between A and palpus obtained with the ***half-sib/full-sib*** and

the ***animal model*** were not congruent and this may be due to the fact that the heritability of palpus is relatively low, resulting in less precise genetic estimates correlations as shown by our simulation study. Phenotypic and genetic correlations between the A and B wing veins are in the same direction. This is probably a result of developmental constraints on the shape of wing cells to ensure efficient flight. Our results are in line with previous studies showing genetic correlation between morphometric traits in honeybee (Collins et al., 1984; Poklukar & Kezić, 1994).

To summarize, wing traits are highly heritable and seem robust to environmental variation. These traits also show divergent pattern between populations and subspecies and are therefore useful for morphological determination. On the contrary, palpus length is less heritable (particularly in Mauritius) and displays higher phenotypic plasticity: it cannot be used to classify specimens in the South-West Indian Ocean.

ACKNOWLEDGMENTS

We thank Olivier Esnault for providing honeybee samples and Meenowa Deodass and Sookar Preeaduth for their help in fieldwork in Mauritius. We also thank Elodie Chapuis and two anonymous reviewers for their comments on the manuscript and Matthew Wolak for his help with **nadiv** package. The authors acknowledge the South Green Platform (http://www.southgreen.fr) for providing computational resources that have contributed to the research results reported within this paper. This study has been financially supported by the CIRAD, the "Conseil Régional de La Réunion" and the European Agricultural Fund for Rural Development (EAFRD). The authors acknowledge the Plant Protection Platform (3P, IBISA) where molecular experiments were conducted.

CONFLICT OF INTEREST

The authors have declared that no conflict of interest exists.

AUTHOR CONTRIBUTIONS

Hélène Jourdan-Pineau: Conceptualization (lead); formal analysis (equal); methodology (lead); supervision (equal); writing – original draft (lead). **Gaëlle Antoine:** Data curation (lead); formal analysis (equal); methodology (equal). **Julien Galataud:** Data curation (equal); formal analysis (equal). **Hélène Delatte:** Conceptualization (equal); funding acquisition (equal); project administration (equal); writing – review and editing (equal). **Christophe Simiand:** Data curation (equal). **Johanna Clémencet:** Conceptualization (equal); formal analysis (equal); funding acquisition (equal); project administration (equal); supervision (equal); writing – review and editing (equal).

DATA AVAILABILITY STATEMENT

Morphological data, genotype, and pedigree are available on Dryad https://doi.org/10.5061/dryad.tx95x69vd.

ORCID

Hélène Jourdan-Pineau https://orcid.org/0000-0002-5118-9873

REFERENCES

Baudry, E., Solignac, M., Garnery, L., Gries, M., Cornuet, J. M., & Koeniger, N. (1998). Relatedness among honeybees (*Apis mellifera*) of a drone congregation. *Proceedings of the Royal Society B-Biological Sciences*, *265*, 2009–2014. https://doi.org/10.1098/rspb.1998.0533

Bérénos, C., Ellis, P. A., Pilkington, J. G., & Pemberton, J. M. (2014). Estimating quantitative genetic parameters in wild populations: A comparison of pedigree and genomic approaches. *Molecular Ecology*, *23*, 3434–3451. https://doi.org/10.1111/mec.12827

Bernstein, R., Plate, M., Hoppe, A., & Bienefeld, K. (2018). Computing inbreeding coefficients and the inverse numerator relationship matrix in large populations of honey bees. *Journal of Animal Breeding and Genetics*, *135*, 323–332. https://doi.org/10.1111/jbg.12347

Bienefeld, K., Ehrhardt, K., & Reinhardt, F. (2007). Genetic evaluation in the honey bee considering queen and worker effects – A BLUP-Animal Model approach. *Apidologie*, *38*, 77–85. https://doi.org/10.1051/apido:2006050

Bienefeld, K., & Pirchner, F. (1990). Heritabilities for several colony traits in the honeybee (*Apis mellifera carnica*). *Apidologie*, *21*, 175–183. https://doi.org/10.1051/apido:19900302

Bohidar, N. R. (1964). Derivation and estimation of variance and covariance components associated with covariance between relatives under sexlinked transmission. *Biometrics*, *20*, 505. https://doi.org/10.2307/2528492

Bolker, B. M. (2008). *Ecological models and data in R*. (pp. 1–396). Ecol. Model. Data R. Princeton University Press.

Brandes, C. (1988). Estimation of heritability of learning behavior in honeybees (*Apis mellifera capensis*). *Behavior Genetics*, *18*, 119–132. https://doi.org/10.1007/BF01067081

Brascamp, E. W., & Bijma, P. (2014). Methods to estimate breeding values in honey bees. *Genetics Selection Evolution*, *46*, 1–15. https://doi.org/10.1186/s12711-014-0053-9

Brascamp, E. W., Willam, A., Boigenzahn, C., Bijma, P., & Veerkamp, R. F. (2016). Heritabilities and genetic correlations for honey yield, gentleness, calmness and swarming behaviour in Austrian honey bees. *Apidologie*, *47*, 739–748.

Brown, G. H. (1969). An empirical study of the distribution of the sample genetic correlation coefficient. *Biometrics*, *25*, 63–72. https://doi.org/10.2307/2528678

Charmantier, A., & Réale, D. (2005). How do misassigned paternities affect the estimation of heritability in the wild? *Molecular Ecology*, *14*, 2839–2850. https://doi.org/10.1111/j.1365-294X.2005.02619.x

Class, B., & Brommer, J. E. (2020). Can dominance genetic variance be ignored in evolutionary quantitative genetic analyses of wild populations? *Evolution (New York)*, *74*(7), 1540–1550. https://doi.org/10.1111/evo.14034

Collins, A. M., Brown, M. A., Rinderer, T. E., Harbo, J. R., & Tucker, K. W. (1987). Heritabilities of honey-bee alarm pheromone production. *Journal of Heredity*, *78*, 29–31. https://doi.org/10.1093/oxfordjournals.jhered.a110302

Collins, A. M., Rinderer, T. E., Harbo, J. R., & Brown, M. A. (1984). Heritabilities and correlations for several characters in the honey bee. *Journal of Heredity*, *75*, 135–140. https://doi.org/10.1093/oxfordjournals.jhered.a109888

Cornuet, J.-M., Fresnaye, J., & Tassencourt, L. (1975). Discrimination Et Classification De Populations D'Abeilles a Partir De Caractères Biométriques. *Apidologie*, *6*, 145–187. https://doi.org/10.1051/apido:19750204

Costa-Maia, F. M., Toledo, V. D. A. A. D., Martins, E. N., Lino-Lourenço, D. A., Sereia, M. J., Oliveira, C. A. L. D., Faquinello, P., & Halak, A. L. (2011). Estimates of covariance components for hygienic behavior in Africanized honeybees (*Apis mellifera*). *Revista Brasileira De Zootecnia*, *40*, 1909–1916. https://doi.org/10.1590/S1516-35982011000900010

Crow, J. F., & Roberts, W. C. (1950). Inbreeding and homozygosis in bees. *Genetics*, *35*, 612–621. https://doi.org/10.1093/genetics/35.6.612

Crozier, R. H. (1970). On the potential for genetic variability in haplodiploidy. *Genetica*, *41*, 551–556. https://doi.org/10.1007/BF00958934

Csilléry, K., Johnson, T., Beraldi, D., Clutton-Brock, T., Coltman, D., Hansson, B., Spong, G., & Pemberton, J. M. (2006). Performance of marker-based relatedness estimators in natural populations of outbred vertebrates. *Genetics*, *173*, 2091–2101. https://doi.org/10.1534/genetics.106.057331

De Villemereuil, P., Gimenez, O., & Doligez, B. (2013). Comparing parent-offspring regression with frequentist and Bayesian animal models to estimate heritability in wild populations: A simulation study for Gaussian and binary traits. *Methods in Ecology and Evolution*, *4*, 260–275. https://doi.org/10.1111/2041-210X.12011

Diniz-Filho, J. A. F., Bueno, O. C., Chaud-Netto, J., & Malaspina, O. (1993). Heritability of the number of ovarioles in honey bee workers (*Apis mellifera* L.) (hym.:Apidae). *Revista Brasileira De Genética*, *16*, 917–921.

Estoup, A., Solignac, M., & Cornuet, J. M. (1994). Precise assessment of the number of patrilines and of genetic relatedness in honeybee colonies. *Proceedings of the Royal Society B-Biological Sciences*, *258*, 1–7.

Faquinello, P., De Arnaut De Toledo, V. A., & Martins, E. N., De Oliveira, C. A. L., Sereia, M. J., Costa-Maia, F. M., & Colla Ruvolo-Takasusuki, M. C. (2011). Parameters for royal jelly production in Africanized honeybees. *Sociobiology*, *57*, 495–510.

Fernando, R. L., & Grossman, M. (1990). Genetic evaluation with autosomal and X-chromosomal inheritance. *TAG. Theoretical and Applied Genetics*, *80*, 75–80. https://doi.org/10.1007/BF00224018

Firth, J. A., Hadfield, J. D., Santure, A. W., Slate, J., & Sheldon, B. C. (2015). The influence of nonrandom extra-pair paternity on heritability estimates derived from wild pedigrees. *Evolution (New York)*, *69*(5), 1336–1344. https://doi.org/10.1111/evo.12649

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

Fjerdingstad, E. J. (2005). Control of body size of Lasius niger ant sexuals - worker interests, genes and environment. *Molecular Ecology*, *14*, (10), 3123–3132. https://doi.org/10.1111/j.1365-294X.2005.02648.x.

Gilmour, A. R., Gogel, B. J., Cullis, B. R., Thompson, R., Butler, D. (2009). *ASReml user guide release 3.0*. VSN Int. Ltd.

Goudie, F., Allsopp, M. H., Beekman, M., Lim, J., & Oldroyd, B. P. (2012). Heritability of worker ovariole number in the cape honey bee *Apis mellifera capensis*. *Insectes Sociaux*, *59*, 351–359. https://doi.org/10.1007/s00040-012-0227-9

Harbo, J. R. (1992). Breeding honey bees (Hymenoptera: Apidae) for more rapid development of larvae and pupae. *Journal of Economic Entomology*, *85*, 2125–2130. https://doi.org/10.1093/jee/85.6.2125

Harpur, B. A., Chernyshova, A., Soltani, A., Tsvetkov, N., Mahjoorighasrodashti, M., Xu, Z., & Zayed, A. (2014). No genetic tradeoffs between hygienic behaviour and individual innate immunity in the honey bee, *Apis mellifera*. *PLoS One*, *9*, e104214. https://doi.org/10.1371/journal.pone.0104214

Holand, A. M., & Steinsland, I. (2016). Is my study system good enough? A case study for identifying maternal effects. *Ecology and Evolution*, *6*(11), 3486–3495. https://doi.org/10.1002/ece3.2124

Koffler, S., de Matos Peixoto Kleinert, A., & Jaffé, R. (2017). Quantitative conservation genetics of wild and managed bees. *Conservation Genetics*, *18*, 689–700. https://doi.org/10.1007/s10592-016-0904-8

Kruuk, L. E. B., & Hadfield, J. D. (2007). How to separate genetic and environmental causes of similarity between relatives. *Journal of Evolutionary Biology*, *20*, 1890–1903. https://doi.org/10.1111/j.1420-9101.2007.01377.x

Laidlaw, H. H., & Page, R. E. (1984). Polyandry in honey bees (*Apis mellifera* L.): Sperm utilization and intracolony genetic relationships. *Genetics*, *108*, 985–997.

Laloi, D., & Pham-Delegue, M. H. (2010). Patriline-level variability in olfactory learning in the honey bee. *Apidologie*, *41*, 436–442. https://doi.org/10.1051/apido/2009080

Le Conte, Y., Bruchou, C., Benhamouda, K., Gauthier, C., & Cornuet, J. M. (1994). Heritability of the queen brood post-capping stage duration in *Apis mellifera mellifera* L. *Apidologie*, *25*, 513–519. https://doi.org/10.1051/apido:19940601

Liu, F. H., & Smith, S. M. (2000). Estimating quantitative genetic parameters haplodiploid organisms. *Heredity (Edinburg)*, *85*, 373–382.

Lush, J. L., & Hubbs, C. L. (1945). *Animal breeding plans*. Iowa State College Press.

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

Melo, A., Malaspina, O., & Diniz-Filho, J. A. F. (1997). Heritability of sting characters in Africanized honeybees. *Journal of Venomous Animals and Toxins*, *3*, 274–279. https://doi.org/10.1590/S0104-79301997000200003

Milne, C. P., & Friars, G. W. (1984). An estimate of the heritabiltty of honeybee pupal weight Complementary lethal factors in two North American populations of the yellow monkey flower. *Journal of Heredity*, 509–510.

Moilanen, A., Sundstroem, L., & Pedersen, J. S. (2004). MATESOFT: A program for deducing parental genotypes and estimating mating system statistics in haplodiploid species. *Molecular Ecology Notes*, *4*, 795–797. https://doi.org/10.1111/j.1471-8286.2004.00779.x

Morimoto, H. (1968). The Use of the Labial Palpus as a Measure of Proboscis Length in Worker Honeybees *Apis mellifera ligustica* and *Apis cerana cerana*. *Journal of Apicultural Research*, *7*, 147–150.

Moritz, R. F. A. (1985). Heritability of the postcapping stage in *Apis mellifera* and its relation to varroatosis resistance. *Journal of Heredity*, *76*, 267–270. https://doi.org/10.1093/oxfordjournals.jhered.a110090

Moritz, R. F. A., & Hillesheim, E. (1985). Inheritance of dominance in honeybees (*Apis mellifera capensis* Esch.). *Behavioral Ecology and Sociobiology*, *17*, 87–89. https://doi.org/10.1007/BF00299434

Moritz, R. F. A., & Klepsch, A. (1985). Estimating heritabilities of worker characters: A new approach using laying workers of the cape honeybee (*Apis mellifera capensis* Esch.). *Apidologie*, *16*, 47–56.

Morrissey, M. B., Wilson, A. J., Pemberton, J. M., & Ferguson, M. M. (2007). A framework for power and sensitivity analyses for quantitative genetic studies of natural populations, and case studies in Soay sheep (*Ovis aries*). *Journal of Evolutionary Biology*, *20*, 2309–2321.

Mostajeran, M., Edriss, M. A., & Basiri, M. R. (2006). Analysis of colony and morphological characters in honey bees (*Apis mellifera meda*). *Pakistan Journal of Biological Sciences*, *9*, 2685–2688. https://doi.org/10.3923/pjbs.2006.2685.2688

Mousseau, T. A., & Roff, D. A. (1987). Natural selection and the heritability of fitness components. *Heredity (Edinburg)*, *59*, 181–197. https://doi.org/10.1038/hdy.1987.113

Oldroyd, B., & Moran, C. (1983). Heritability of worker characters in the honeybee (*Apis mellifera*). *Australian Journal of Biological Sciences*, *36*, 323–332. https://doi.org/10.1071/BI9830323

Oldroyd, B., Rinderer, T., & Buco, S. (1991). Heritability of morphological characters used to distinguish European and Africanized honeybees. *TAG. Theoretical and Applied Genetics*, *82*, 499–504. https://doi.org/10.1007/BF00588605

Pemberton, J. M. (2008). Wild pedigrees: The way forward. *Proceedings of the Royal Society B-Biological Sciences*, *275*, 613–621. https://doi.org/10.1098/rspb.2007.1531

Pernal, S. F., Sewalem, A., & Melathopoulos, A. P. (2012). Breeding for hygienic behaviour in honeybees (*Apis mellifera*) using free-mated nucleus colonies. *Apidologie*, *43*, 403–416. https://doi.org/10.1007/s13592-011-0105-x

Poklukar, J., & Kezić, N. (1994). Estimation of heritability of some characteristics of hind legs and wings of honeybee workers (*Apis mellifera carnica* Polm). *Apidologie*, *25*, 3–11.

Postma, E. (2014). Four decades of estimating heritabilities in wild vertebrate populations: Improved methods, more data, better estimates. In A. Charmantier, D. Garant, & L. E. B. Kruuk (Eds.), *Quantitative genetics in the wild* (p. 33). Oxford University Press.

Rinderer, T. E. (1977). Measuring the heritability of characters of honeybees. *Journal of Apicultural Research*, *16*, 95–98. https://doi.org/10.1080/00218839.1977.11099867

Rinderer, T. E., Collins, A. M., & Brown, M. A. (1983). Heritabilities and correlations of the honey bee: Response to *Nosema apis*, longevity, and alarm response to isopentyl acetate. *Apidologie*, *14*, 79–85.

Ruttner, F. (1988). *Biogeography and taxonomy of honeybees*. Springer.

Ruttner, F., Tassencourt, L., & Louveaux, J. (1978). Biometrical-statistical analysis of the geographic variability of *Apis mellifera* L. I. Material and methods. *Apidologie*, *9*, 363–381. https://doi.org/10.1051/apido:19780408

Sheehan, M. J., Choo, J., & Tibbetts, E. A. (2017). Heritable variation in colour patterns mediating individual recognition. *Royal Society Open Science*, *4*, 161008. https://doi.org/10.1098/rsos.161008

Solignac, M., Vautrin, D., Loiseau, A., Mougel, F., Baudry, E., Estoup, A., Garnery, L., Haberl, M., & Cornuet, J.-M. (2003). Five hundred and fifty microsatellite markers for the study of the honeybee (*Apis mellifera* L.) genome. *Molecular Ecology Notes*, *3*, 307–311. https://doi.org/10.1046/j.1471-8286.2003.00436.x

Szabo, T. I., & Lefkovitch, P. L. (1992). Heritability of colour patterns in a closed population of honeybees. *Apidologie*, *23*, 151–159. https://doi.org/10.1051/apido:19920208

Techer, M. A., Clémencet, J., Simiand, C., Preeaduth, S., Azali, H. A., Reynaud, B., & Hélène, D. (2017). Large-scale mitochondrial DNA analysis of native honey bee *Apis mellifera* populations reveals a new African subgroup private to the South West Indian Ocean islands. *BMC Genetics*, *18*(1), 1–21. https://doi.org/10.1186/s12863-017-0520-8

Techer, M. A., Clémencet, J., Simiand, C., Turpin, P., Garnery, L., Reynaud, B., & Delatte, H. (2017). Genetic diversity and differentiation among insular honey bee populations in the southwest Indian Ocean likely reflect old geographical isolation and modern introductions. *PLoS One*, *12*, 1–26. https://doi.org/10.1371/journal.pone.0189234

Wang, J. (2006). Informativeness of genetic markers for pairwise relationship and relatedness inference. *Theoretical Population Biology*, *70*, 300–321. https://doi.org/10.1016/j.tpb.2005.11.003

Wilson, A. J., Réale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L. E. B., & Nussey, D. H. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology*, *79*, 13–26. https://doi.org/10.1111/j.1365-2656.2009.01639.x

Wolak, M. E. (2012). nadiv: An R package to create relatedness matrices for estimating non-additive genetic variances in animal models. *Methods in Ecology and Evolution*, *3*, 792–796.

Wolak, M. E., & Keller, L. F. (2014). Dominance genetic variance and inbreeding in natural populations. In A. Charmantier, D. Garant, & L. E. B. Kruuk (Eds.), *Quantitative genetics in the wild* (pp. 104–127). Oxford University Press.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Jourdan-Pineau H, Antoine G, Galataud J, Delatte H, Simiand C, Clémencet J. Estimating heritability in honeybees: Comparison of three major methods based on empirical and simulated datasets. *Ecol Evol*. 2021;00:1–12. https://doi.org/10.1002/ece3.7389