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

Hélène Jourdan-Pineau[1], Gaëlle Antoine[2], Julien Galataud[3], Hélène DELATTE[2], Christophe Simiand[4], and Johanna Clémencet[5]

[1]CIRAD Biological Systems Department
[2]CIRAD Département Systèmes biologiques
[3]Université de la Réunion Faculté des Sciences et Technologies
[4]CIRAD Departement Systemes biologiques
[5]Université de la Réunion UFR Sciences de l'Homme et de l'Environnement

June 15, 2020

Abstract

The level of the genetic contribution to phenotypic variation (namely the heritability) determines 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, sib-analyses are frequently used with three main methods: an historical model using the average colony relatedness, a half-sibs/full-sibs 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 853 workers in total. All individuals were genotyped to reconstitute the pedigree, and phenotypic traits were measured: the proboscis- and wing-associated lengths. We also simulated phenotypic datasets with varying levels of heritability, common environmental effect and genetic correlation between traits. The simulation approach showed that the average colony relatedness was highly biased in presence of common environmental effect whereas the half-sibs/full-sibs 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-related traits had high heritabilities, allowing the use of those morphometric characters to discriminate between populations. On the contrary, the palpus length (associated to proboscis) was more sensitive to environmental factors. Finally, significant genetic correlations among measured traits indicate that they do not evolve independently.

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 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 response of the trait, h^2 the heritability (*sensus stricto*) 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 complex polyandrous species. First, the morphological, physiological and behavioural differences between castes preclude the use of straightforward parent-offspring regression approach (with the exception of some traits, e.g. Szabo and Lefkovitch 1992) and require sibling analyses. A noticeable exception is the Cape honeybee, *Apis mellifera capensis* , in which laying workers produce female workers, allowing parent-offspring regression to be performed (Moritz & Hillesheim, 1985; Moritz & Klepsch, 1985; Brandes, 1988; Le Conte*et al.* , 1994). More importantly, the general framework of quantitative genetics is based on diploid organisms, reproducing randomly in populations of infinite size. If non-compliance with the last two assumptions 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 the queensto 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 $r = \frac{1}{4}$) (e.g. Rinderer et al. 1983; Collins et al. 1984, 1987; Harbo 1992). Heritability is then estimated from the intra-colony 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 and Page 1984 or Milne and 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 (Oldroyd & Moran, 1983; Oldroyd *et al.* , 1991; Diniz-Filho *et al.* , 1993; Poklukar & Kezić, 1994; Melo *et al.* , 1997; Goudie *et al.* , 2012). 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 as well as by the environment provided by the mother.

When paternal lineages within colonies can be reconstituted (thanks to controlled fertilization or to 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 males-drones) is the natural situation in the honeybee and this leads to different formulae, as stressed out by Fjerdingstad (Fjerdingstad, 2005; Laloi & Pham-Delegue, 2010; Harpur *et al.* , 2014).

With molecular genotyping, the additive genetic relationship matrix among individuals (derived from the pedigree) can be reconstituted 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.* , 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 (Crow & Roberts, 1950; Bohidar, 1964; Crozier, 1970; Fernando & Grossman, 1990). This method is implemented in the R package **nadiv** (Wolak, 2012), which has been applied to wasps but never to honeybees (Sheehan *et al.* , 2017).

However, estimating heritability on each trait separately does not reflect the actual evolutionary processes and is a poor predictor of response to selection. Indeed, traits do not evolve independently from each other, due to linkage or pleiotropy (one loci influencing several traits) (Lynch & Walsh, 1998). Therefore 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 1000 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 (morphology, behaviour, life history traits...). Koffler et al (Koffler *et al.* , 2017) reviewed all published estimates in honeybee (and other bee, ant and wasp species) and they showed that heritability estimates were affected by the trait type (as 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) but not by the analytical methods used (among which sibling analyses and parent-offspring regression) or the sample size. To our view, this result, contrasting with previously published studies (*e.g.* Postma 2014), could be due to biases inherent to meta-analysis: incomplete information on the design or the analyses conducted, datasets diverging on several tested factors (resulting in confounded effects). However a few studies compared different methods in the same honeybee dataset and found that those methods yielded heritability estimates in the same range (Moritz, 1985; Moritz & Klepsch, 1985; Brandes, 1988). Genetic correlation estimates are scarce (14 studies only in honeybee) and half 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, design...) based on their accuracy of estimates and on their statistical power to detect heritability and genetic correlation (Morrissey *et al.* , 2007; De Villemereuil *et al.* , 2013; Holand & Steinsland, 2016).

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. Those traits (associated with mouthpart and forewing morphometry) are used to discriminate between species or populations (Cornuet *et al.* , 1975; Ruttner*et al.* , 1978; Ruttner, 1988). We also generated simulated phenotypic datasets with varying levels of heritability, genetic correlation and common environment effect.

Material and methods

Empirical dataset

Honeybee workers were sampled in Mauritius and La Reunion, two islands of the Mascareigne archipelago situated in the South West Indian Ocean. In this area, honeybee subspecies is *Apis melliferaunicolor* belonging to the African lineage with recent import of the European lineage (Techer *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). Those colonies were equally distributed in Mauritius and in La Reunion. All colonies of Mauritius were sampled in October 2014, whereas colonies of La Reunion were sampled from 2011 to 2018 (Annex 1). Individuals were kept in 95 ° alcohol at -80 ° C until they were processed.

Patrilines were determined by genotyping eight microsatellite markers on DNA extracted from the femur as described in Delatte et al. 2005. 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 *et al.* , 2017). PCR reactions were 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 30s at 94°C, 30s at 52°C, 45s 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. Thus, 424 individuals were kept from Mauritius; and 429 individuals from La Reunion (Annex 2). Patriline reconstruction was carried out on the basis of allelic frequencies previously estimated for those insular populations (Techer *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.

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 (Carl Zeiss AG, n.). To study the variation in size of the mouthparts, we measured the length of the long segment of the right palpus (Annex 3) 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 (CI) 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 *et al.* , 1978; Ruttner, 1988). Repeated measures of the palpus length and cubital cell of 36 individuals with two montages and two measurements for each of these montages were realized. Measurement errors were mainly due to slide mounting and it was almost null for the cubital cell (A, B and CI) and less than 2% for the palpus length. Therefore experimental noise would not artificially increase the residual variance and bias heritability.

All measured traits significantly differed between colonies and between islands (except for CI) (Annex 4). Morphological measurements were globally larger in La Reunion than in Mauritius. In addition, a negative phenotypic correlation between the A and B veins and a positive correlation between B vein and palpus length were observed (Annex 5).

Quantitative genetics analysis methods

As mentioned above, three commonly used methods based on sibling analysis were compared for their performances in estimating heritabilities and genetic correlations: 1) a simple linear model considering a colony effect and an average colony relatedness between individuals (Oldroyd & Moran, 1983); 2) a nested model of patrilines in the colonies equivalent to a half-sib/full-sib approach (Fjerdingstad, 2005) and 3) 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 (Liu & Smith, 2000; Fjerdingstad, 2005). Thus only the broad sense heritability is estimated:$H^2 = \frac{V_G}{V_P}$. In addition, those methods do not take into account the epistatic variance, which 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 a rough estimation of the number of efficient matings (m) is needed. Heritability is then calculated as follows:

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

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

Genetic correlation (r_G) is given by:

$r_{G=}\frac{\text{cov}_{\text{col}_{t1t2}}}{\sqrt{V_{\text{col}_{t1}} * V_{\text{col}_{t2}}}}$ with $\text{cov}_{\text{col}_{t1t2}}$ the covariance between traits t1 and t2 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:

$H^2 = \frac{2 \times V_{\text{pat}}}{V_{\text{pat}} + V_{\text{CE}} + V_R}$(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$V_{\text{pat}} = \frac{1}{2}V_A + \frac{1}{2}V_D$.

The genetic correlation is estimated by:

$r_{G=}\frac{\text{cov}_{\text{pat}_{t1t2}}}{\sqrt{V_{\text{pat}_{t1}} * V_{\text{pat}_{t2}}}}$ with $\text{cov}_{\text{pat}_{t1t2}}$ the covariance between traits t1 and t2 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 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. Here diploid females are coded as the heterogametic sex whereas haploid males are the homogametic sex. A "colony" 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:

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

Genetic correlations between two traits (t1 and t2) are calculated based on bivariate animal models according to the formula:

$r_G = \frac{\text{cov}_{G_{t1t2}}}{\sqrt{V_{A_{t1}} * V_{A_{t2}}}}$with cov_G the genetic covariance between traits t1 and t2.

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

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.

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, CI 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). No fixed effects were fitted in the statistical models.

Using the animal model , we also tested if the additive, common environment and residual variances differed between islands by adding a random island effect. Log-likelihood ratio tests between nested models (with or without random island effect) allowed to test if variances were significantly different between islands.

Quantitative genetics analyses of simulated datasets

To test the performances of the three methods in estimating heritability and genetic correlations, simulations have been performed.

Heritability. Following the approach of Morrissey et al. (Morrissey *et al.* , 2007), we first simulated a dataset of individual phenotypic values such that$Y_i = \mu + A_i + EC_k + \varepsilon_i$ 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); EC_k , the common environment effect (normally distributed with V_{EC}variance) and ϵ_ι , the residual variation (normally distributed with V_R variance). The genetic values A_i were computed according to the simulated pedigree and V_A using the **grfx**function of the **nadiv** package (Wolak, 2012) .

In each scenario studied, the heritability (h^2) and common environment (EC^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, 1000 phenotypic datasets were simulated based on our experimental pedigree or on a simulated pedigree. This simulated pedigree was built to be of similar size and design to our experimental pedigree: it is composed of 840 phenotyped individuals (against 853 in our real pedigree) 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. This simulated pedigree corresponds to a balanced design.

Each simulated phenotypic data set was analysed with the three previously described methods (average colony relatedness, half-sib/full-sib and animal model) to estimate the heritability and significance of the "genetic" effect (corresponding to the additive genetic matrix).

To further compare the half-sib/full-sib and the animal model , we also generated a modified pedigree from our dataset, where 3 queens of 3 different colonies were considered as full-sisters, therefore adding relatedness relationship between those 3 colonies. This modified pedigree was used to generate phenotypic datasets (with heritability and common environment varying as indicated above).

The statistical performances of each method were evaluated considering the accuracy of estimations (according to the mean and root mean square error of estimates for one simulated value) and the percentage of simulated datasets with a significant genetic effect.

Genetic correlation. We simulated two correlated traits with heritabilities values of 0.1, 0.3 and 0.5 (we skipped the null value since there could not be genetic correlation between traits that are not heritable) and genetic correlation of 0, 0.1, 0.3 and 0.5 (according to the genetic correlation estimates published in honeybee (Koffler*et al.* , 2017)) using the **grfx** function of the**nadiv** R-package. A common environment effect of 0.3 was added for the half-sib/full-sib and the animal model methods whereas it was not included for the average colony relatedness method. The statistical performances of the three methods were evaluated as above (precision and power).

RESULTS

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 EC^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 (and p-values) (Table 2). According to the animal model , the genetic correlations were significant between A and B veins (r_G = - 0.49) and also between A vein and the palpus length (r_G = 0.21). None of the correlation were significant using the half-sib/full-sib method, whereas only the cubital index/palpus correlation was significant using average colony relatedness method.

Using the animal model , variance estimates were not significantly different between islands except for palpus length (Table 3). Indeed palpus length heritability was lower in Mauritius (H^2= 0.18) than in La Reunion (H^2 = 0.46). This difference was mostly due to an almost 10-times higher common environment variance in Mauritius.

Simulations

Analyses carried out with the three methods with simulated datasets based on the simulated pedigree provided estimates close to the expected values in the absence of common environment effect (Fig. 1, upper part). When a common environment effect (EC^2) was added, the average colony relatedness method largely overestimated heritabilities, resulting in h^2 values greater than 1, whereas the half-sib/full-sib and the animal model methods provided similarly reliable estimates (though with slight upward bias of around 8% for $EC^2 = 0.5$, Fig 1). However, in this context, the animal model delivers the lowest estimation error (as shown by the smallest RMSE). Using a balanced pedigree decreased the RMSE associated with h^2 of the half-sib/full-sib method such that the precision was similar to the one obtained with the animal model (Annex 6).

For the three methods (and the 3 tested pedigrees), the rate of false positive (probability to obtain a significant genetic effect whereas$h^2 = 0$) was low, except for the average colony relatedness method in presence of a common environment effect (Fig. 1. right side). The statistical power was low for heritabilities of 0.1 but satisfying for moderate to high heritabilities (h^2 [?] 0.3).

The half-sib/full-sib and the animal model methods were not affected by the presence of related queens in the simulated pedigree and yielded similar results as mentioned above for our empirical pedigree, except that the RMSE increased (especially with the half-sib/full-sib method) (Annex 7).

Concerning genetic correlations, the average colony relatedness method gave very unreliable estimates (Fig. 2). 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 (Fig 3 a). The statistical power was low when genetic correlation was lower than 0.5 (except when the two traits were highly heritable, fig 3. b). Finally, the animal model is a bit more precise than the half-sib/full-sib method. Note that 2 to 4% of the models failed to converge, independently of the method (data not shown).

DISCUSSION

On 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 and Moran 1983). 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 (Csillery *et al.* , 2006; Wang, 2006; Pemberton, 2008). Here, the number of patrilines estimated for each colony fluctuates according to the number microsatellite markers used: expectedly, with 7 markers instead of 8, the number of patrilines estimates per colony was lower. Similarly, heritability estimates were lower with 7 microsatellite markers (Annex 8). 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 (Charmantier & Reale, 2005; Berenos *et al.* , 2014; Firth *et al.* , 2015) and have little impact on general interpretations.

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 for example (Oldroyd & Moran, 1983; Moritz & Hillesheim, 1985)).

Both the half-sib/full-sib and the animal model provided reliable heritability estimates with sufficient power (in particular with intermediate to high heritabilities), either on our experimental pedigree or on the simulated pedigree. However, the half-sib/full-sib method had lower precision than the animal model except when

using a balanced design, in which case both methods yielded similar RMSE. This greater performance of the animal model had already been acknowledged and is based on the greater flexibility of the approach that takes all types of relationships into account between individuals, maximizing statistical power and providing more accurate estimates (Kruuk, 2004). Because of this flexibility, the animal model method is the most suitable for studying the heritability of traits measured on a (semi-) natural population where crossings are not controlled (Wilson *et al.* , 2010).

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's. This assumption was reasonable since mating occurs during nuptial flight in congregation area where thousands of drones gather, from all colonies in the surroundings (Baudry *et al.* , 1998). In such situation, the copulation of the queen with one drone related to herself or with two drones related to each other's 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. In such situation, the half-sib/full-sib method, which ignores such additional relationship, should 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 (Liu & 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 estimates of the best quality (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 & Pirchner, 1990; Bienefeld *et al.* , 2007) 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 as well as a dominance component (Liu & Smith, 2000), which could explain the high heritabilities for the characters studied. However, 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 honeybee in La Reunion and Mauritius, the effective population size seems to be large (Techer *et al.* , 2017) which suggest a limited effect of dominance on the genetic variance. One meta-analysis showed that the dominance variance is only important for domesticated species and is generally low (0.15, (Wolak & Keller, 2014)).

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 and 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 environmental effect. Estimates were very 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 1000 individuals, (Brown, 1969; Lynch & Walsh, 1998; De Villemereuil*et al.* , 2013). In honeybee, colony sizes are very large, allowing to easily sample a large number of individuals, hence the only limitation is the ability to phenotype and genotype all sampled individuals.

Our results on the empirical dataset are consistent with the simulation outputs: the very high heritability estimates provided by the average colony relatedness method are in agreement with the presence of common environment effect (EC^2<0.5 for the palpus length). According to the simulation results, we discuss only the estimates provided by the animal model . The high heritability estimates obtained for all traits are consistent with the literature ((Moritz & Klepsch, 1985; Oldroyd *et al.* , 1991; Poklukar & Kezić, 1994; Mostajeran *et al.* , 2006) reviewed in (Koffler *et al.* , 2017)). Morphological traits are known to display higher heritabilities than fitness-related traits owing to Fisher's fundamental theorem (Fisher, 1958). In addition, 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 in Mauritius than in La Reunion due an almost 10-times larger common environment variance. Accordingly, the proboscis is the morphological character showing the largest geographic variability (Ruttner *et al.* , 1978; Ruttner, 1988) supporting a great phenotypic plasticity. In addition, colonies came from more diverse environments in 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 the palpus. Those genetic correlations are in the same direction as the phenotypic correlations. Hence the evolution of the cubital cell shape is constraint by the genetic correlation and this is probably a result of developmental constraint on the wing to ensure efficient flight. The genetic (and phenotypic) correlation between palpus length and A vein is probably a mere allometric relationship. 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, providing good resolving power to discriminate subspecies or ecotypes. On the contrary, palpus length is less heritable (particularly in Mauritius) and display higher phenotypic plasticity: it cannot be used to classify specimens in the South West Indian Ocean.

Data accessibility:

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

Bibliographic references

Baudry, E., Solignac, M., Garnery, L., Gries, M., Cornuet, J.M. & Koeniger, N. 1998. Relatedness among honeybees (Apis mellifera) of a drone congregation. *Proc. R. Soc. B Biol. Sci.* **265** : 2009–2014.

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. *Mol. Ecol.***23** : 3434–3451.

Bernstein, R., Plate, M., Hoppe, A. & Bienefeld, K. 2018. Computing inbreeding coefficients and the inverse numerator relationship matrix in large populations of honey bees. *J. Anim. Breed. Genet.***135** : 323–332.

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.

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

Bohidar, N.R. 1964. Derivation and Estimation of Variance and Covariance Components Associated with Covariance between Relatives under Sexlinked Transmission. *Biometrics* **20** : 505.

Brandes, C. 1988. Estimation of heritability of learning behavior in honeybees (Apis mellifera capensis). *Behav. Genet.* **18** : 119–132.

Brascamp, E.W. & Bijma, P. 2014. Methods to estimate breeding values in honey bees. *Genet. Sel. Evol.* **46** : 1–15.

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.

Charmantier, A. & Réale, D. 2005. How do misassigned paternities affect the estimation of heritability in the wild? *Mol. Ecol.***14** : 2839–2850.

Collins, A.M., Brown, M.A., Rinderer, T.E., Harbo, J.R. & Tucker, K.W. 1987. Heritabilities of honey-bee alarm pheromone production. *J. Hered.* **78** : 29–31.

Collins, A.M., Rinderer, T.E., Harbo, J.R. & Brown, M.A. 1984. Heritabilities and correlations for several characters in the honey bee. *J. Hered.* **75** : 135–140.

Cornuet, J.-M., Fresnaye, J. & Tassencourt, L. 1975. Discrimination Et Classification De Populations D'Abeilles a Partir De Caractères Biométriques. *Apidologie* **6** : 145–187.

Costa-Maia, F.M., de Toledo, V. de A.A., Martins, E.N., Lino-Lourenço, D.A., Sereia, M.J., de Oliveira, C.A.L., *et al.* 2011. Estimates of covariance components for hygienic behavior in Africanized honeybees (Apis mellifera). *Rev. Bras. Zootec.* **40** : 1909–1916.

Crow, J.F. & Roberts, W.C. 1950. Inbreeding and homozygosis in bees. *Genetics* **35** : 612–621.

Crozier, R.H. 1970. On the potential for genetic variability in haplo-diploidy. *Genetica* **41** : 551–556.

Csilléry, K., Johnson, T., Beraldi, D., Clutton-Brock, T., Coltman, D., Hansson, B., *et al.* 2006. Performance of marker-based relatedness estimators in natural populations of outbred vertebrates. *Genetics* **173** : 2091–2101.

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 Ecol. Evol.* **4** : 260–275.

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). *Rev. Bras. Genet.* **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. *Proc. R. Soc. B Biol. Sci.* **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., *et al.* 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. *Theor. Appl. Genet.* **80** : 75–80.

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 (N. Y).***69** : 1336–1344.

Fisher, R.A. 1958. *The genetical theory of natural selection* , 2nd edn. Dover Publications, New York.

Fjerdingstad, E.J. 2005. Control of body size of Lasius niger ant sexuals - Worker interests, genes and environment. *Mol. Ecol.***14** : 3123–3132.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Thompson, R., Butler, D. & others. 2009. ASReml user guide release 3.0. *VSN Int. Ltd, Hemel Hempstead, UK* .

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 Soc.* **59** : 351–359.

Harbo, J.R. 1992. Breeding Honey Bees (Hymenoptera: Apidae) for More Rapid Development of Larvae and Pupae. *J. Econ. Entomol.***85** : 2125–2130.

Harpur, B.A., Chernyshova, A., Soltani, A., Tsvetkov, N., Mahjoorighasrodashti, M., Xu, Z., *et al.* 2014. No genetic tradeoffs between hygienic behaviour and individual innate immunity in the honey bee, Apis mellifera. *PLoS One* **9** .

Holand, A.M. & Steinsland, I. 2016. Is my study system good enough? A case study for identifying maternal effects. *Ecol. Evol.* 1–10.

Koffler, S., de Matos Peixoto Kleinert, A. & Jaffé, R. 2017. Quantitative conservation genetics of wild and managed bees.*Conserv. Genet.* **18** : 689–700.

Kruuk, L.E.B. 2004. Estimating genetic parameters in natural populations using the "animal model." *Philos. Trans. R. Soc. B Biol. Sci.***359** : 873–890.

Laidlaw, H.H. & Page, R.E. 1984. Polyandry in Honey Bees (Apis mellifera L.): Sperm Utilization and Intracolony Genetic Relationships.*Genetics* **108** : 985–97.

Laloi, D. & Pham-Delegue, M.H. 2010. Patriline-level variability in olfactory learning in the honey bee. *Apidologie* **41** : 436–442.

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.

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

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

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

Melo, A., Malaspina, O. & Diniz-Filho, J.A.F. 1997. Heritability of sting characters in Africanized honeybees. *J. Venom. Anim. Toxins***3** : 274–279.

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. *J. Hered.* 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. *Mol. Ecol. Notes* **4** : 795–797.

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. *J. Apic. Res.* **7** : 147–150.

Moritz, R.F.A. 1985. Heritability of the postcapping stage in Apis mellifera and its relation to varroatosis resistance. *J. Hered.***76** : 267–270.

Moritz, R.F.A. & Hillesheim, E. 1985. Inheritance of dominance in honeybees (Apis mellifera capensis Esch.). *Behav. Ecol. Sociobiol.* **17** : 87–89.

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). *J. Evol. Biol.* **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 J. Biol. Sci.* **9** : 2685–2688.

Oldroyd, B. & Moran, C. 1983. Heritability of worker characters in the honeybee (apis mellifera). *Aust. J. Biol. Sci.* **36** : 323–332.

Oldroyd, B., Rinderer, T. & Buco, S. 1991. Heritability of morphological characters used to distinguish European and Africanized honeybees. *Theor. Appl. Genet.* **82** : 499–504.

Pemberton, J.M. 2008. Wild pedigrees: The way forward. *Proc. R. Soc. B Biol. Sci.* **275** : 613–621.

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.

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: *Quantitative genetics in the wild* , p. 33. Oxford University Press Oxford, UK.

Rinderer, T.E. 1977. Measuring the Heritability of Characters of Honeybees. *J. Apic. Res.* **16** : 95–98.

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. Berlin, Germany.

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.

Sheehan, M.J., Choo, J. & Tibbetts, E.A. 2017. Heritable variation in colour patterns mediating individual recognition. *R. Soc. Open Sci.* **4** : 161008.

Solignac, M., Vautrin, D., Loiseau, A., Mougel, F., Baudry, E., Estoup, A., *et al.* 2003. Five hundred and fifty microsatellite markers for the study of the honeybee (Apis mellifera L.) genome. *Mol. Ecol. Notes* **3** : 307–311.

Szabo, T.I. & Lefkovitch, P.L. 1992. Heritability of colour patterns in a closed population of honeybees. *Apidologie* **23** : 151–159.

Techer, M.A., Clémencet, J., Simiand, C., Turpin, P., Garnery, L., Reynaud, B., *et al.* 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.

Wang, J. 2006. Informativeness of genetic markers for pairwise relationship and relatedness inference. *Theor. Popul. Biol.***70** : 300–321.

Wilson, A.J., Réale, D., Clements, M.N., Morrissey, M.M., Postma, E., Walling, C.A., *et al.* 2010. An ecologist's guide to the animal model. *J. Anim. Ecol.* **79** : 13–26.

Wolak, M.E. 2012. nadiv: an R package to create relatedness matrices for estimating non-additive genetic variances in animal models.*Methods Ecol. Evol.* **3** : 792–796.

Wolak, M.E. & Keller, L.F. 2014. Dominance genetic variance and inbreeding in natural populations. In: *Quantitative Genetics in the Wild* , pp. 104–127. Oxford University Press Oxford, UK.

Table 1: Estimated genetic parameters of the measured traits estimated (the cubital veins A and B, the cubital index (CI) 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. 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. All genetic effects were statistically significant.

Method	Trait	V_{col}	V_{col}	V_R	H^2	SE (H^2)
Average colony relatedness	A	3.37E-04	3.37E-04	1.10E-03	0.89	0.34
	B	1.18E-04	1.18E-04	4.78E-04	0.75	0.30
	CI	1.80E-02	1.80E-02	1.08E-01	0.54	0.24
	palpus	4.84E-03	4.84E-03	2.02E-03	2.68	0.38
HS/FS		V_{pat}	V_{CE}	V_R	H^2	SE (H^2)
	A	2.99E-04	3.64E-04	8.14E-04	0.41	0.09
	B	1.66E-04	9.11E-05	2.91E-04	0.61	0.09
	CI	3.05E-02	1.71E-02	7.63E-02	0.49	0.09
	palpus	8.59E-04	4.89E-03	1.14E-03	0.25	0.09
Animal model		V_G	V_{CE}	V_R	H^2	SE (H^2)
	A	5.99E-04	2.14E-04	6.64E-04	0.41	0.09
	B	3.32E-04	8.12E-06	2.08E-04	0.61	0.09
	CI	6.09E-02	1.90E-03	6.11E-02	0.49	0.09
	palpus	1.72E-03	4.46E-03	7.06E-04	0.25	0.09

Table 2: Genetic correlation between pairs of measured traits (the cubital veins A and B, the cubital index (CI) 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). Significant genetic correlations (p-value < 0.05) are highlighted in bold.

Method	A-B	A-palpus	B-palpus	CI-palpus
Average colony relatedness	0.17	0.72	0.05	**0.42**
Half-sib/full-sib	-0.57	-0.46	-0.49	-0.07
Animal model	**-0.49**	**0.21**	0.02	0.06

Table 3: Estimated genetic parameters of the measured traits (the cubital veins A and B, the cubital index (CI) and the length of the long segment of the right palpus) estimated using an animal model, with distinct parameters on each island (as random effect). We presented values for genetic variance (V_G), variance associated with common environment (V_{EC}) and residual variance (V_R), heritability (H^2), and its standard error (SE) for each island. P-values associated with the island random effect.

Trait	Mauritius	Mauritius	Mauritius	Mauritius	Mauritius	La Reunion	La Reunion	La Reunion	La Reuni
	V_G	V_{EC}	V_R	H^2	SE (H^2)	V_G	V_{EC}	V_R	H^2
A	5.22E-04	5.19E-04	6.75E-04	0.30	0.12	6.65E-04	1.46E-08	6.62E-04	0.50
B	3.32E-04	4.00E-05	2.21E-04	0.56	0.13	3.18E-04	3.20E-08	2.01E-04	0.61
CI	6.19E-02	4.61E-03	6.54E-02	0.47	0.11	5.84E-02	1.74E-07	5.78E-02	0.50
palpus	1.95E-03	8.12E-03	4.94E-04	0.18	0.10	1.47E-03	8.31E-04	8.99E-04	0.46

Figure 1: Performance of heritability estimates evaluated from simulated datasets based on our experimental design, depending on the model used (average colony relatedness (star), half-sib/full-sib (triangle) and animal model (circle), from light grey to black). Left side: mean and RMSE of heritability estimates (h^2) (shapes and error bars, respectively, y-axis). Right side: power to determine genetic effect (from 0 to 100%, y-axis), as a function of simulated heritabilities (x-axis) and common environmental effects (EC^2, vertical panels).

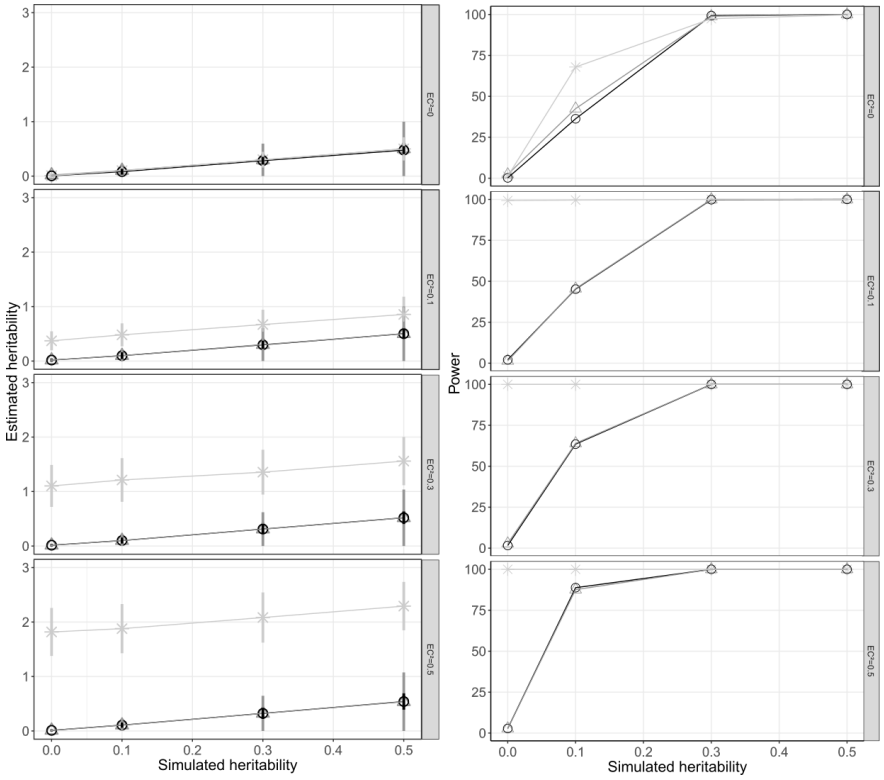

Figure 2: Performance of genetic correlation estimates evaluated from simulated datasets based on our experimental design, using the average colony relatedness method. Left side: mean and RMSE of genetic correlation estimates (r_G) (circles and error bars, respectively, y-axis). Right side: power to determine genetic correlation, as a function of simulated r_G (x-axis) and simulated heritabilities of the two traits (h^2 of t1 and t2, horizontal and vertical panels).

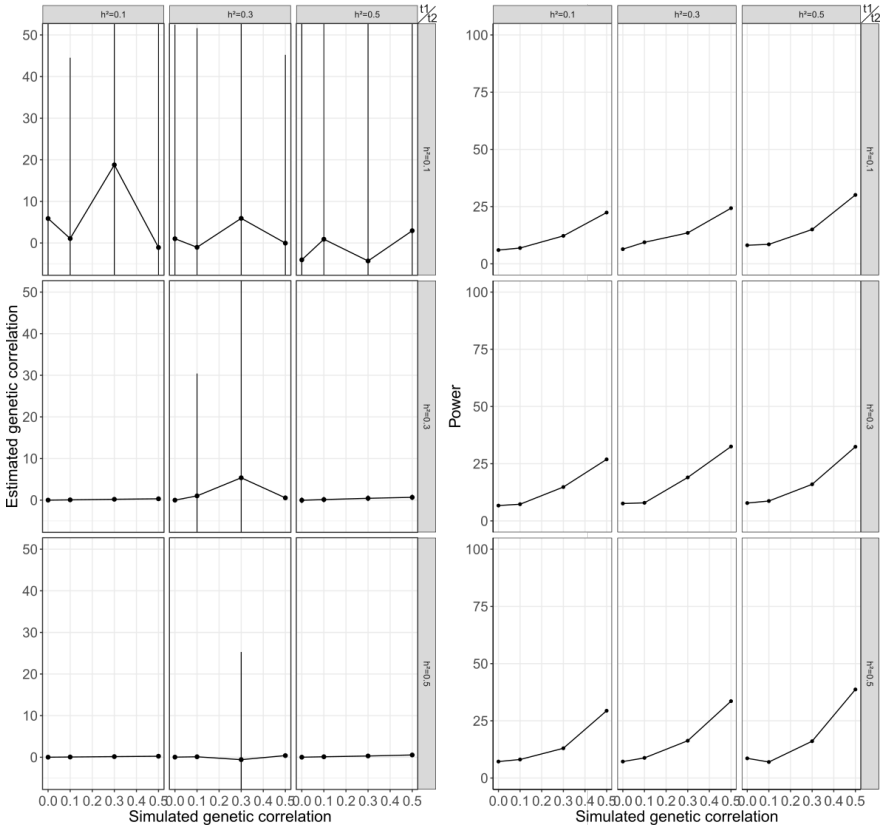

Figure 3: Performance of genetic correlation estimates evaluated from simulated datasets based on our experimental design, depending on the model used (half-sib/full-sib and animal model, grey and black, respectively). Left side: mean and RMSE of genetic correlation estimates (r_G) (circles and error bars, respectively, y-axis). Right side: power to determine genetic correlation, as a function of simulated r_G (x-axis) and simulated heritabilities of the two traits (h^2 of t1 and t2, horizontal and vertical panels).