

# Hygienic behaviour in honeybees: a comparison of two recording methods and estimation of genetic parameters

Elena FACCHINI<sup>1</sup>, Piter BIJMA<sup>2</sup>, Giulio PAGNACCO<sup>1</sup>, Rita RIZZI<sup>1</sup>, Evert Willem BRASCAMP<sup>2</sup>

<sup>1</sup>Department of Veterinary Medicine, University of Milano, via G. Celoria 10, 20133, Milan, Italy

<sup>2</sup>Animal Breeding and Genomics, Wageningen University & Research, PO Box 338, 6700AH, Wageningen, The Netherlands

Received 20 July 2018 – Revised 4 December 2018 – Accepted 14 December 2018

**Abstract** – Hygienic behaviour (HB) in honeybees reflects social immunity against diseases and parasites. Young bees showing HB detect, uncap, and remove infested brood from a colony. We developed a new variant of freeze-killed brood (FKB\*) test to optimise the duration of the HB test, the costs, and safety for the operator. In 2016, we performed a comparison between traditional FKB and FKB\* on 25 unselected and unrelated colonies in the apiary of the University of Milano. To estimate repeatability and heritability, in 2017 and 2018, FKB\* was used to phenotype, respectively, 56 and 95 colonies twice, in the context of a breeding programme. FKB\* took less time and required a smaller amount of liquid nitrogen. The two methods showed a correlation between colony effects of 0.93, indicating that they measure the same trait. For single records, the phenotypic correlation between both methods was 0.64. Estimated heritability and repeatability for single records HB were 0.23 and 0.24, respectively, whilst heritability for the average HB value of two records was 0.37.

honeybee / hygienic behaviour / freeze-killed brood test / repeatability / heritability

## 1. INTRODUCTION

Hygienic behaviour (HB) is known as a behavioural response of honeybee workers to the spreading of infections in the colony, conferring resistance against diseases and parasites that affect, inter alia, the honeybee brood (Rothenbuhler, 1964; Wilson-Rich et al. 2009). Indeed, hygienic workers can detect the presence of an infected larva or pupa and react by uncapping the wax cover of a brood cell, if the cell was sealed, and by removing the diseased individual. Hygienic behaviour evolved as a general mechanism of resistance to the brood

pathogens including *Paenibacillus larvae* (the causative agent of American foulbrood), *Ascosphaera apis* (the causative agent of chalkbrood), and the parasitic mite *Varroa destructor* (Gilliam et al. 1983; Spivak and Reuter 1998; Harbo and Harris 1999; Spivak and Reuter 2001). Hygienic behaviour has a genetic basis, and it is a heritable trait (Rothenbuhler 1964; Moritz 1988; Kefuss et al. 1996). Therefore, it soon became an object for selective breeding programmes worldwide (Spivak 1996; Spivak et al. 2009; Büchler et al. 2010; Büchler et al. 2013). Although the relevance of genomics for bee improvement programmes is likely to increase, phenotyping remains essential. Therefore, the recording of this trait by reliable and cheap field assays is crucial to estimate breeding values.

Currently, HB is recorded by assessing the dead brood removal rate of a colony. There are two principal methods described in the

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s13592-018-0627-6>) contains supplementary material, which is available to authorized users.

Corresponding author: E. Facchini,  
elena.facchini@unimi.it  
Handling editor: Cédric Alaux

literature: the mechanical killing of brood by using an entomological needle, known as the “pin-test”, and the thermal killing by using low temperature through liquid nitrogen or a freezer (Newton and Ostasiewski 1986; Momot and Rothenbuhler 1971; Spivak and Reuter 1998). The basic idea of the two methods is to sacrifice a determined area of sealed brood in the hive and to record how much the worker bees clean that area by removing dead larvae from it in a fixed time window, usually 24 h for the thermal killing method. The pin-test is used more frequently for its cheapness and simplicity, but it damages the pupae under the wax cap, with possible haemolymph leakage, which could affect the test result as it boosts the cleaning stimulus (Spivak and Downey 1998; Panasiuk et al. 2008). The thermal killing procedures have good discriminatory ability but are more expensive regarding equipment because of the need for liquid nitrogen or an extra trip to the apiary when a freezer is used (Espinosa Montaña et al. 2008; Büchler et al., 2010; Kefuss et al., 2015).

We developed a modification of the standard methodology described in Spivak and Reuter (1998) in the expectation that it would result in a more practical field assay and perform better. In essence, FKB\* is the combination of FKB for liquid nitrogen usage and the freezer method for cutting out the brood disc. We expected a better-defined test area, lower time requirement, and lower nitrogen use. In 2016, a field trial was carried out to compare the modified method (FKB\*) with the standard methodology (FKB). For an improved method to be used in a breeding programme, it is important to have estimates of the repeatability and heritability of the trait. In 2017 and 2018, therefore, HB was recorded using the FKB\* method on a testing population in the context of a breeding programme.

Here, we first describe the FKB\* methodology for recording HB. Next, we present the results of a field comparison between FKB and FKB\* and present estimates of heritability and repeatability for HB recorded with FKB\* in another trial.

## 2. MATERIAL AND METHODS

### 2.1. Comparison of the two methods

#### 2.1.1. Location and study colonies

The comparison was conducted at the apiary of the Veterinary Faculty of Milan, located in Lodi, Italy, during spring/summer of 2016. A total of 25 colonies were included in the field test. All colonies had good health status and were headed by naturally mated queens bought from different Italian breeders. Information on the pedigree of the colonies was not available; therefore, we assumed that these colonies were unrelated. Each colony was kept in a Dadant-Blatt hive box with ten frames.

#### 2.1.2. FKB

The FKB method was described in Spivak AND Reuter (1998). The materials needed for this test are a tube, liquid nitrogen, a camera (optional), and safety equipment. The method consists in extracting a capped brood comb from the colony to be tested and finding a suitable portion of brood to maximise the number of capped cells covered by the tube. Then, the tube is twisted on the comb until the mid-rib of the frame. Ca. 300 ml of liquid nitrogen is poured into the tube to freeze-kill the brood delimited by the tube. Once the liquid nitrogen is evaporated and the tube is thawed, the tube is extracted from the comb and a photo is taken. The comb is marked to be easily distinguishable and repositioned in the colony of origin. After 24 h, the same comb is taken out from the hive and the treated area of brood is photographed for the count of removed brood.

#### 2.1.3. FKB\*

The key feature of FKB\* is that the brood area to be tested is cut out and frozen through immersion in liquid nitrogen in an insulating bowl, rather than on the comb directly. Therefore, FKB\* is only compatible with wax foundation brood combs.

The material needed for this method is a thin metal tube with a sharp end, liquid nitrogen, a

camera (optional), safety equipment (insulating gloves and safety goggles), tweezers, water-based ink marker, and an insulating polystyrene box. The tube must have a diameter that allows it to pass between the iron wires of the frame (ca. 6–8 cm, depending on the frame type). A comb is taken from the hive to be tested, a suitable portion of capped brood (of any age) preferably on both sides of the comb is found, and the tube is twisted in the brood in order to pass through the comb and to cut out a brood disc. A good practise is to mark the brood disc with the water-based ink marker to track back its original position and orientation. The liquid nitrogen is poured into the polystyrene box. The brood disc is taken out from the tube and dipped in liquid nitrogen. After ca. 2 min, the brood disc is fished out using tweezers, allowed to thaw for 3 min, and then repositioned in the brood comb. A photo is taken of both sides to permit the count of sealed cells at time zero. The comb is marked on the top and placed back in the hive. After 24 h, the same comb is taken out and a photo of each side is taken for further analysis of dead brood removal. If many colonies need to be tested, we suggest to place inside the insulated box cardboard walls in order to subset the internal part of the vessel. In this way, it is possible to keep track of the brood discs that are dipped together in the liquid nitrogen. See Figure S1 of supplementary materials for an illustrated step-by-step guide.

### 2.1.4. Experimental design

The two methods were applied at the same time to the same comb in each colony, where the location of both was random. We chose this approach to minimise the potential variation due to the comb and to the distribution of worker bees in the colony. The tests were repeated six times during the spring/summer of 2016. During the experiment, the composition of the tested colonies sometimes changed due to swarming and to hive condition (presence of capped brood). Therefore, not all 25 colonies were phenotyped for every replicate, but the number of replicates per colony ranged from 2 to 5. In total, 74 observations for each method were available for the subsequent statistical analysis. The time spent per test and colony was recorded.

### 2.1.5. Photo analysis and HB scoring

The counting of dead brood removal was performed by analysing the pictures that were taken for each tested area at time zero and after 24 h. Image analysis was performed with the help of the counter tool of the software ImageJ (Schneider et al. 2012). HB was recorded as the proportion of removed dead larvae in 24 h:

$$HB = 1 - \frac{\text{sealed cells } T24}{\text{sealed cells } T0}$$

HB was scored in the most conservative way, i.e., if the cell was only partly uncapped or if it was uncapped but the dead larva was only partially removed, the cell was considered sealed.

### 2.1.6. Statistical analysis

The objective was to compare the HB results of the two methods in terms of average HB values, repeatability of both methods, and correlation between the two methods. In addition, to quantify the benefit of repeatedly recording HB, we derived the accuracy of the mean HB score of a colony, as a function of the number of records.

A paired-sample *t* test was conducted to compare the average HB value of a colony recorded with the two methods.

To estimate repeatability, a univariate approach was adopted for each method. The following mixed model was fitted to the data:

$$y_{ijkl} = \mu + B_i + T_j + C_{ki} + e_{ijkl},$$

where  $\mu$  represents the overall mean,  $B$  is the fixed effect of the  $i$ th breeder of origin ( $i = 1, 5$ ), and  $T$  term represents the fixed effect of the  $j$ th replicate ( $j = 1, 6$ );  $C$  represents the random effect the  $k$ th colony within the  $i$ th breeder (the size of  $k$  varies between breeders of origin from 2 to 7), and  $e$  represents the random error term of the  $l$ th observation, where  $l$  varies between colonies from 2 to 5 due to swarming and capped brood availability.

Interest is in the  $C$  term, which represents the HB effect of the colony including all genetic and permanent environmental effects, whereas the  $e$  term represents the temporary environmental

effect and measurement error (e.g., due to the location of the tube on the comb). The effect of the breeder of origin was included as fixed term to avoid the inflation of the  $C$  variance component due to the differences between the genetic sources of the colonies. The colony effect and the error were assumed to follow a normal distribution with means zero and variances  $\sigma_C^2$  and  $\sigma_e^2$ , respectively.

Repeatability ( $r$ ), which is the correlation between repeated records on the same colony (Falconer and Mackay 1996), was estimated for each method by the following formula:

$$r = \frac{\sigma_C^2}{\sigma_C^2 + \sigma_e^2} = \frac{\sigma_C^2}{\sigma_P^2}$$

where  $\sigma_P^2$  is the phenotypic variance for a single records. The repeatability also measures the reliability of the estimated  $C$  value of a colony, based on a single record (Falconer and Mackay 1996).

In addition, a bivariate analysis was conducted to estimate the correlation between the  $C$  terms of the two HB scores, applying the model mentioned above. Note that the  $C$  term represents the colony effect of interest, so we measured the similarity of both traits by  $r_C$  rather than the phenotypic correlation  $r_P$ . If  $r_C$  is close to one, both methods essentially represent the same trait, apart from temporary measurement error ( $e$ ).

Furthermore, we calculated the accuracy of each method as a function of the number of records. Our interest is in  $C$ , and the phenotype is the mean of  $n$  repeated records,  $\bar{y}$ . Thus, the accuracy is defined as the correlation between  $C$  and  $\bar{y}$ . The relationship between the accuracy and  $n$  reveals the benefit of repeatedly recording HB. The accuracy for each method was calculated by the following formula (see Appendix for the derivation):

$$a = \frac{\sigma_C}{\sqrt{\sigma_C^2 + \frac{\sigma_e^2}{n}}}$$

The trends of the accuracy of each method and the phenotypic correlation between the two methods were plotted as a function of the number of records.

Statistical analyses were performed using the computing environment R (R Core Team 2015). Mixed models were fitted using the R package lme4 (Bates et al. 2015).

## 2.2. Estimation of heritability and repeatability

### 2.2.1. Colonies and phenotyping

Heritability and repeatability were estimated only for the FKB\* method, from data collected in 2017 and 2018. FKB\* was used to phenotype a cohort of 151 colonies made available by Melyos, an Italian bee-breeding and beekeeping company. The colonies were kept in two apiaries, near Zelo Buon Persico, Lodi, Italy, during 2017 and in one apiary in 2018 located in Lesmo, Monza, Italy. The tested group was composed of colonies headed by groups of sister queens with known pedigree, all naturally mated at an isolated mating station hosting one paternal line which was different for the two groups tested in the 2 years. Each colony was managed in the context of a breeding programme, therefore in the most standardised way. The colonies were phenotyped twice for HB during productive season of 2017 and of 2018.

### 2.2.2. Statistical analysis

For the analysis, estimates of the genetic relationships between groups of workers and queen in colonies are required, and also those with the groups of drone-producing queens with which queens are mated. We used the methods of Brascamp and Bijma (2014) to estimate these relationships. The pedigree file was built following the procedure described in Brascamp et al. (2016). To estimate heritability and repeatability, the statistical package ASReml and the pin function of the nadiv package were used in the computing environment R (Butler 2009; Wolak 2012; R Core Team 2015). Only the genetic effect of the workers was included, as the paucity of data did not allow us to simultaneously estimate the queen and the worker's effect. Following Brascamp et al. (2018), we used the additive genetic variance of worker groups to calculate phenotypic variance.

First, we fitted the overall average of HB for each colony, using the following mixed model:

$$y_{ijk} = \mu + ApY_{ij} + A_{wijk} + e_{ijk}$$

where  $\mu$  represents the overall mean,  $ApY$  is the fixed effect of the combination of the  $i$ th apiary ( $i = 1, 2, 3$ ) and the  $j$ th year ( $j = 1, 2$ );  $A_w$  represents the random genetic effect of the  $k$ th colony where the number of colonies per apiary varies between 17, 39, and 95, 151 in total ( $k = 1, 151$ ) and  $e$  represents the random error term. This model allowed us to estimate the heritability of the mean value of two HB records measured with FKB\*.

Secondly, we fitted a repeatability animal model:

$$y_{ijklm} = \mu + ApTY_{ijk} + A_{wijkl} + pe_l + e_{ijklm}$$

where  $\mu$  represents the overall mean,  $ApTY$  is the fixed effect term representing the combination of the  $i$ th apiary ( $i = 1, 2, 3$ ) and  $j$ th recording time ( $j = 1, 2$ ) and the  $k$ th year of observation ( $k = 1, 2$ );  $A_w$  represents the random genetic effect of the  $l$ th colony (151 colonies in total),  $pe$  represents the random permanent environmental effect of the  $l$ th colony ( $l = 1, 151$ ), and  $e$  represents the random error term. This model allowed us to estimate both heritability and repeatability for HB.

In order to compare the two models, we inspected the accuracies of estimated breeding values of colonies for both methods.

### 3. RESULTS AND DISCUSSION

#### 3.1. Comparison between the two methods

The aim of the first part of this study was to compare only FKB\* with the standard FKB. However, other methods are available to measure dead brood removal ability of a bee colony (Büchler et al. 2013). Table 1 reports some practical aspects of recording HB with FKB\* compared to FKB. FKB\* was found to take less time in the field, since no evaporation time is required. On the other hand, FKB\* requires the analysis of four instead of two pictures for HB calculation. Concerning materials, FKB\* required less liquid nitrogen, because it is possible to freeze many

brood discs at the same time repeatedly using the same liquid nitrogen since it is kept in an insulating box. FKB\* is safer because it requires less handling of liquid nitrogen, reducing the chance to get burned by accidental spilling. Moreover, FKB\* requires only one tube (or a few if many colonies have to be recorded simultaneously and more than one operator is performing the test).

A visual comparison of the two methods is shown in Figure 1. It can be noticed that FKB\* produced clear borders of the killed area on both sides of the brood frame with no evidence of collateral brood damages (Figure 1c, d; blue circles), giving complete control over the amount of killed brood. Figure 1 also shows that FKB, which is carried out only on one side, is capable of killing the brood on the other side of the treated area. Indeed, in Figure 1d, a large patch of removed brood is visible in correspondence of the FKB which was carried out on the other side of the comb.

An empirical feature was that the bees, regardless of their HB score, tended to clean perfectly all the brood that was physically and irremediably damaged by the tube. We observed this phenomenon in both methods. As described by Spivak and Downey (1998) and Panasiuk et al. (2008), we also noticed that the mechanical injury may trigger the stimulus of dead removal. Therefore, in our HB estimation, we did not consider the cells on the circle in both methods.

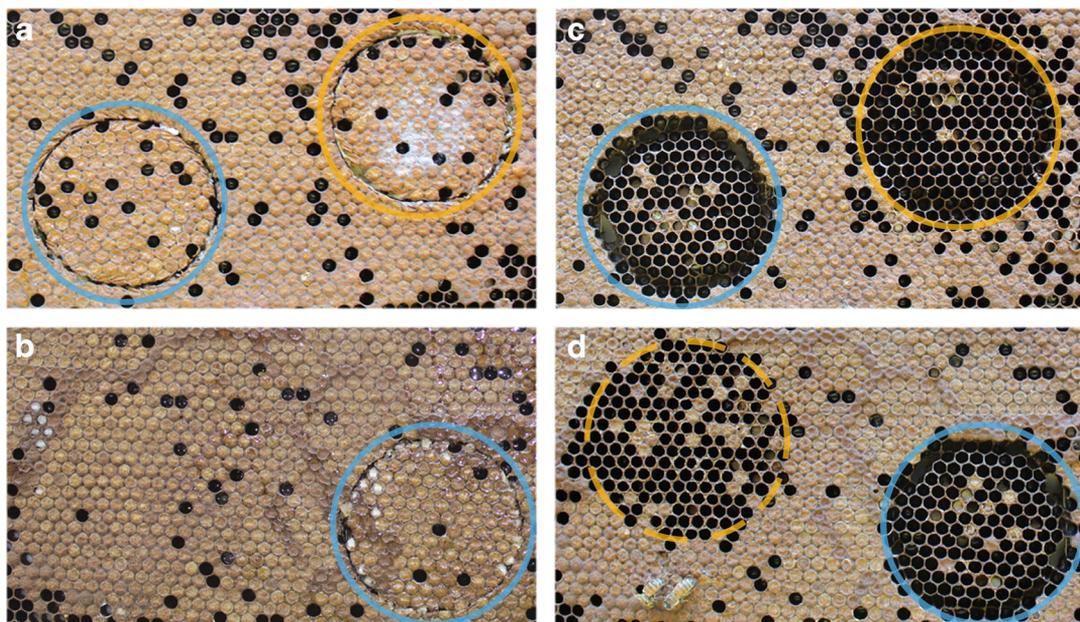
Comparing the results of the two recording methods, we found that HB measured with FKB ( $m = 0.59$ ,  $sd = 0.21$ ) was significantly lower than with FKB\* ( $m = 0.70$ ,  $sd = 0.17$ ) with an estimated mean difference of  $-0.11$  ( $t = -4.80$ ,  $df = 24$ ,  $P < 0.001$ ). This result indicates that on average HB score is higher if measured with FKB\* compared to FKB. The consistent lower score produced by FKB test could be due to the fact that the standard test is killing a broader area compared to the one considered for the calculation of HB. Therefore, they may spend time cleaning outside the tested area lowering the HB result. These collateral damages are represented by the surroundings of the tube which can be killed by cold fumes whilst liquid nitrogen is evaporating and by the correspondent brood on the other side of the comb which can be killed by the deep freezing treatment.

**Table I.** Practical aspects of the FKB and FKB\* methods

	FKB	FKB*
Liquid nitrogen	~ 3 colonies/l	~ 8 colonies/l
Time	15–20 min/colony	7–10 min/colony
Tubes (cylinders)	2 for each colony	1 for all colonies
Tested area	1 side	2 sides
Photo analysis	2/colony	4/colony

Table II reports the estimated variance components from the univariate model and correlations from the bivariate model. Results of the univariate model show that almost half of the total variance is explained by the effect of the colony, in both methods. Repeatability was slightly higher for FKB\* (0.48) than that for FKB (0.42). Both estimates are close to the value of 45.5% reported by Bigio et al. (2013), who repeated the FKB test 10 times on a cohort of 19 unrelated and unselected colonies.

The correlation of the colony effects was very high (0.93), which implies that the two recording methods essentially measure the same trait. Indeed, the correlation of the colony effects comprises all genetic and permanent environmental effects of HB. The phenotypic correlation between single observations with the two methods was clearly lower (0.63), which indicates that the correlation of the temporary measurement errors (0.42) is much lower than the correlation of colony effects.



**Figure 1.** Visual comparison of the two methods. The pictures show the two sides (namely A and B) of a tested comb where the two methods were performed simultaneously. Blue circles surround FKB\*-tested brood disc, and orange circles surround FKB-tested brood area. **a** Side A of the comb at time zero, which shows the brood disc that was cut, frozen, and repositioned for FKB\* (blue circle) and only frozen on the comb for FKB (orange circle). **b** Side B of the comb at time zero, which shows the same portion of comb in picture A but on the other side. **c** Side A of the same comb after 24 h from the test. **d** Side B of the same comb after 24 h from the test. The large patch of removed brood corresponds to the treated area of FKB carried out on side A (orange dashed circle).

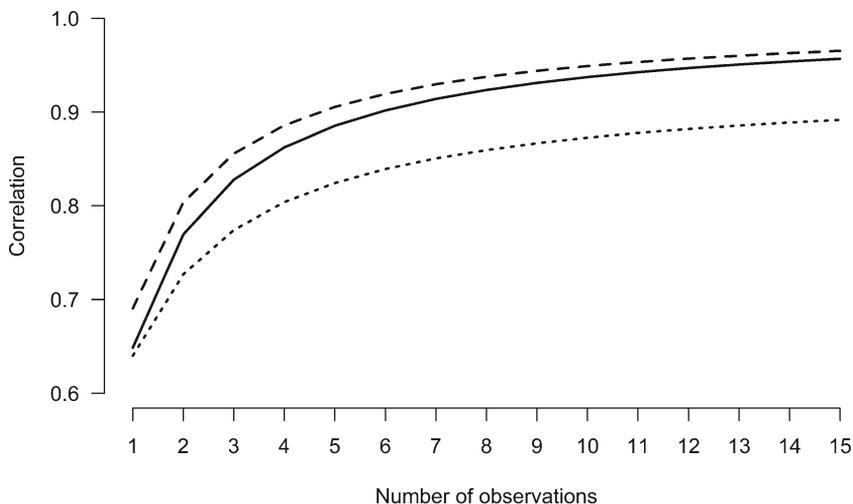
**Table II.** Estimated variance components for HB recorded with the standard method (FKB) and the variant method (FKB\*). Variances (Var) of random colony effect ( $C$ ), residual ( $e$ ), and total ( $P$ ) estimated with univariate model are used to derive the repeatability ( $r$ ) for each method. Phenotypic correlation ( $r_P$ ), correlation of colony effects ( $r_C$ ), and correlation of error term ( $r_e$ ) were derived from variance and covariance components estimated with the bivariate model. Approximate standard errors are reported in brackets

	FKB	FKB*
Var ( $C$ )	0.016 (0.008)	0.012 (0.005)
Var ( $e$ )	0.022 (0.005)	0.013 (0.003)
Var ( $P$ )	0.038 (0.008)	0.024 (0.005)
$r$	0.42 (0.15)	0.48 (0.13)
$r_P$	0.64 (0.09)	
$r_C$	0.93 (0.13)	
$r_e$	0.42 (0.12)	

The phenotypic correlation in Figure 2 shows the similarity between the two methods as a function of the number of records. Repeating the test increases the similarity between the two recording methods, and with many replicates, the phenotypic correlation asymptotes to a maximum equal to

the correlation of the colony effects ( $r_c = 0.93$ ). Therefore, if the test is repeated many times on a colony, the probability to assess the true merit of a colony increases, regardless of the recording method. This can also be seen from the trend of the accuracy for each method shown in Figure 2. The accuracy represents the correlation between the mean of the phenotype measured  $n$  times and the true effect of the colony, i.e., the permanent component of the trait for each recording method. The accuracy increases strongly between 1 and 4 observations. These values are directly linked to repeatability (Appendix 1). For each method, repeating the test at least twice is highly advisable for a more accurate estimate of the HB level of a colony, as illustrated in Figure 2.

The estimates for the environmental effects for each method are represented by the residual variances in Table II. The residual variance for FKB\* (0.013) is almost halved compared to FKB (0.022). Moreover, the correlation between the environmental effects between the two methods (0.42) indicates that temporary variation in the two recording methods is similar but not identical. The lower environmental variance of FKB\* compared to FKB suggests that FKB\* could be successful in eliminating unwanted sources of



**Figure 2.** Correlations for mean of repeated HB records on a colony as a function of the number of observations. Dotted line: phenotypic correlation between FKB\* and FKB calculated with parameters from bivariate model; solid line: accuracy for FKB calculated with parameters from the univariate model; dashed line: accuracy for FKB\* calculated with parameters from the univariate model.

environmental variation. An example could be the collateral killing that occurs with FKB due to the lack of a clear border of the killed area on the other side of the comb (Figure 1c).

### 3.2. Heritability and repeatability

Table III shows the estimates for the variance components and the resulting heritabilities and repeatability of HB recorded with FKB\*. As expected, heritability for the average HB score of two records ( $0.37 \pm 0.25$ ) was higher than the one estimated with the repeatability model ( $0.23 \pm 0.16$ ), but it is smaller compared to the value of  $0.65 \pm 0.61$  reported by Harbo and Harris (1999) and to the values of 0.56 and 0.57 recently published by Guarna et al. (2017). The higher heritability is explained by the fact that in the average model, the dependent variable was the average of two HB measures. Therefore, the total resulting variance was smaller (0.018) than for single records (0.029).

The estimated permanent environmental variance was close to zero ( $2.7 \times 10^{-4}$ ). Therefore, the repeatability estimate was near to heritability (0.24, Table III).

To compare the two models, we computed accuracies of estimated breeding values for each model which appeared to be very similar

**Table III.** Estimated genetic parameters for hygienic behaviour. Variances (Var) of genetic effect for the average of workers ( $w$ ), permanent environmental effect ( $pe$ ), residual ( $e$ ), and phenotypic ( $P$ ). Derived from these are estimates of heritability ( $h^2$ ) and repeatability ( $r$ ).  $\bar{r}_{\hat{A},A}$  is the average accuracy of breeding values. Approximate standard errors are reported in brackets

	Average model	Repeatability model
Var ( $w$ )	0.007 (0.005)	0.007 (0.005)
Var ( $pe$ )	–	0.0003 (0.0039)
Var ( $e$ )	0.011 (0.004)	0.022 (0.003)
Var ( $P$ )	0.018 (0.002)	0.029 (0.003)
$h^2$	0.37 (0.25)	0.23 (0.16)
$r$	–	0.24 (0.09)
$\bar{r}_{\hat{A},A}$	0.50	0.50

suggesting that there is in principle no benefit of a repeatability model over an average model.

### 4. CONCLUSION

FKB and FKB\* are two ways to measure the same trait, i.e., the dead removal ability of a bee colony. We did not investigate in this study the correlation of HB with the biological trait of interest as previous reports showed that HB can be used as a proxy to select colonies for resistance to main brood pathogens (Spivak and Reuter, 2001; Panasiuk et al. 2008; Panasiuk et al. 2014). FKB\* requires less time and liquid nitrogen and has a smaller measurement error, resulting in a slightly higher repeatability. To accurately measure HB, the test should be repeated at least twice. Heritability for the average HB score of two FKB\* recordings was 0.37, indicating good prospects for genetic improvement of HB. Based on accuracies of estimated breeding values (EBVs), there was no benefit of using a repeatability model over the use of a model for the average of two HB score.

### AUTHORS' CONTRIBUTION

EF and RR conceived, planned, and carried out the experiments. GP and RR contributed in the design and supervision of the project. EF, PB, RR, and EWB contributed to the analysis and interpretation of the results. EF drafted the manuscript and designed the figures and tables. PB, RR, and EWB provided critical feedback on the manuscript and participated in the revision of it. All authors read and approved the final manuscript.

### FUNDING INFORMATION

This project was funded by Regione Lombardia, Fondo Europeo Agricolo per lo Sviluppo Rurale (FEASR), Programma Sviluppo Rurale (PSR) 2014-2020, Operazione 16.2.01.

### COMPLIANCE WITH ETHICAL STANDARDS

**Conflict of interest** The authors declare that they have no conflict of interest.

## APPENDIX

In this appendix, we derive the accuracy as the correlation between the colony effect ( $C$ ) and the average HB value measured on a colony with  $n$  observations.

$$a = \text{Corr} \left( C, \bar{y} \right)$$

Applying the definition of correlation, it follows that:

$$a = \frac{\text{Cov} \left( C, \bar{y} \right)}{\sqrt{\text{Var} (C) * \text{Var} \left( \bar{y} \right)}}$$

Considering that  $\bar{y} = \mu + C + e/n$ , where  $n$  indicates the number of observations and assuming no interaction between the genotype and the environment, it follows that

$$a_n = \frac{\sigma_C^2}{\sqrt{\sigma_C^2 \times \left( \sigma_C^2 + \frac{\sigma_e^2}{n} \right)}}$$

Division of numerator and denominator by  $\sigma_C$  gives

$$a = \frac{\sigma_C}{\sqrt{\sigma_C^2 + \frac{\sigma_e^2}{n}}}$$

## OPEN ACCESS

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

**Comportement hygiénique des abeilles domestiques: comparaison de deux méthodes d'enregistrement et estimation de paramètres génétiques**

**Abeille domestique / comportement hygiénique / test de couvain tué par congélation / répétabilité / héritabilité**

**Hygieneverhalten bei Honigbienen: Ein Vergleich von zwei Erfassungsmethoden und Abschätzung genetischer Parameter**

**Honigbiene / Hygieneverhalten / Entfernungstest gefrorener Brut / Wiederholbarkeit / Vererbbarkeit**

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

- Bigio, G., Schürch, R., Ratnieks, F. L. W. (2013) Hygienic behavior in honey bees (Hymenoptera: Apidae): effects of brood, food, and time of the year. *J. Econ. Entomol* 106(6), 2280–2285
- Brascamp, E. W., Bijma, P. (2014) Methods to estimate breeding values in honey bees. *Genet. Sel. Evol* 46, 53
- 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
- Brascamp, E.W., Willam, A., Boigenzahn, C. Bijma, P., Veerkamp, R. F. (2018) Correction to: heritabilities and genetic correlations for honey yield, gentleness, calmness and swarming behaviour in Austrian honey bees. *Apidologie* 49: 462–463 <https://doi.org/10.1007/s13592-018-0573-3>
- Büchler, R., Berg S., Le Conte Y. (2010) Breeding for resistance to Varroa destructor in Europe. *Apidologie* 41, 393–408
- Büchler, R., Andonov S., Bienefeld K., Costa C., Hatijna F., et al. (2013) Standard methods for rearing and selection of *Apis mellifera* queens. *J. Apicult. Res* 52(1): 1–30
- Butler, D. (2009) Asreml: asreml () fits the linear mixed model. R package version 3.0. [www.vsni.co.uk](http://www.vsni.co.uk)
- Bates, D., Maechler M., Bolker B., Walker S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw* 67(1), 1–48
- Espinosa Montaña, L. G., Guzmán Novoa, E., Sánchez Albarrán, A., Montaldo, H. H., Correa Benítez, A. (2008) Comparative study of three assays to evaluate hygienic behavior in honey bee (*Apis mellifera* L.) colonies. *Vet. Mexico* 39(1), 39–54
- Falconer, D. S., Mackay, T. F. C. (1996) Introduction to quantitative genetics. Ed 4. Longmans Green, Harlow, Essex, UK
- Gilliam, M., Taber, S., Richardson, G. V. (1983) Hygienic behaviour of honey bees in relation to chalkbrood disease. *Apidologie* 14, 29–39
- Guarna, M. M., Hoover, S. E., Huxter, E., Higo, H., Moon, K. M., et al. (2017) Peptide biomarkers used for the selective breeding of a complex polygenic trait in honey bees. *Sci. Rep.* 7 (8381), 1–10

- Harbo, J. R., Harris, J. W. (1999) Heritability in honey bees (Hymenoptera: Apidae) of characteristics associated with resistance to *Varroa jacobsoni* (Mesostigmata: Varroidae). *J. Econom. Entomol* 92, 261–265
- Kefuss, J., Taber, S., Vanpoucke, J., Rey, F. (1996) A practical method to test for disease resistance in honey bees. *Am. Bee J* 136, 31–32
- Kefuss, J., Vanpoucke, J., Bolt, M., Kefuss, C. (2015) Selection for resistance to *Varroa destructor* under commercial beekeeping conditions. *J. Apicult. Res* 54(5), 563–576
- Momot, J.P., Rothenbuhler, W.C. (1971) Behaviour genetics of nest cleaning in honeybees. VI. Interactions of age and genotype of bees and nectar flow. *J. Apicult. Res* 10, 11–21
- Moritz, R. (1988) A reevaluation of the two-locus model hygienic behavior in honey bees, *Apis mellifera*. *L. J. Hered* 79, 257–262
- Newton, D.C., Ostasiewski, N.J. Jr. (1986) A simplified bioassay for behavioral resistance to American foulbrood in honey bees (*Apis mellifera* L.). *Am. Bee J* 126, 278–281
- Panasiuk, B., Skowronek, W., Bienkowska, M. (2008) Influence of genotype and method of brood killing on brood removal rate in honey bee. *J. Apicult. Sci* 52(2), 55–65
- R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Panasiuk, B., Bienkowska M., Gerula D., Węgrzynowicz P. (2014) Susceptibility of bee larvae to chalkbrood in relation to hygienic behaviour of worker bees in colonies of chosen races of honey bee (*Apis mellifera*). *J. Apicul. Sci* 58(1): 119–126
- Rothenbuhler, W. C. (1964) Behavior genetics of nest cleaning in honey bee. I. Responses of four inbred lines to disease killed brood. *Anim. Behav* 12, 578–583
- Schneider, C. A.; Rasband, W. S., Eliceiri, K. W. (2012) NIH image to ImageJ: 25 years of image analysis. *Nat. Methods* 9(7), 671–675
- Spivak, M. (1996) Honey bee hygienic behaviour and defense against *Varroa jacobsoni*. *Apidologie* 27, 245–260
- Spivak, M., Reuter, G. S. (1998) Honey bee hygienic behaviour. *Am. Bee J* 138, 283
- Spivak, M., Downey, D. L. (1998) Field assays for hygienic behaviour in honey bees (Hymenoptera: Apidae). *J. Econ. Entomol* 91, 64–70
- Spivak, M., Reuter, G. (2001) Resistance to American foulbrood disease by honey bee colonies *Apis mellifera* bred for hygienic behaviour. *Apidologie* 32(6), 555–565
- Spivak, M., Reuter, G.S., Lee, K., Ranum, B. (2009) The future of the MN hygienic stock of bees is in good hands! *Am. Bee J.* 149, 965–967
- Wilson-Rich, N., Spivak, M., Fefferman, N. H., Starks, P. T. (2009) Genetic, individual, and group facilitation of disease resistance in insect societies. *Annu. Rev. Entomol* 54, 405–423
- Wolak, M.E. (2012) nadv: an R package to create relatedness matrices for estimating non-additive genetic variances in animal models. *Methods Ecol. Evol* 3(5), 792–796