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

Elena Facchini[1], Piter Bijma[2], Giulio Pagnacco[1], Rita Rizzi[1], Evert W. Brascamp[2]

[1]*Department of Veterinary Medicine, University of Milano, via G. Celoria 10, 20133, Milan, Italy.*

[2]*Wageningen University & Research, Animal Breeding and Genomics, PO Box 338, 6700AH, Wageningen, The Netherlands*

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 Freeze Killed Brood test (FKB*) 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 program. 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, while heritability for the average HB value of two records was 0.37.

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 (Rothenbuler, 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 (Rothenbuler 1964, Moritz, 1988; Kefuss et al., 1996). Therefore, it soon became an object for selective breeding programs worldwide (Spivak 1996; Spivak et al., 2009; Büchler et al., 2010). 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 literature: 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 hours for the thermal killing method. The pin-test is used more frequently for its cheapness and simplicity, but it has been reported to have poorer repeatability and discriminatory ability (Spivak and Downey, 1998; Espinosa Montaño et al. 2008; Panasiuk et al. 2008). Moreover, the pin-test damages the pupae, 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 a 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 Montaño et al., 2008; Büchler, 2010; Kefuss, 2015).

We developed a modification of the standard methodology described in Spivak & Reuter (1998) in the expectation that it would result in a more practical field assay and perform better. In essence 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 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 Italian breeders. Information on the pedigree of the colonies was not available. Each colony was kept in a Dadant-Blatt hive box with ten frames.

2.1.2.FKB

The FKB method was described in Spivak & 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 maximize the number of capped cells covered by the tube. Then, the tube is twisted on the comb till 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 hours 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.

The material needed for this method is a tube, liquid nitrogen, a camera (optional), safety equipment, tweezers 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 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 practice is to mark the brood disc 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 minutes, the brood disc is fished out using tweezers, allowed to thaw for 3 minutes, 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 hours the same comb is taken out and a photo of each side is taken for further analysis of dead brood

removal.

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 spring/summer 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 phenotypes 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.

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 hours. 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 hours:

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

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 of 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 μ 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 σ^2_C and σ^2_e, 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 σ_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 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 152 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 single but different 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 two years. Each colony was managed in the context of a breeding program, therefore in the most standardized 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 workers 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:

$$y_{ij} = \mu + ApY_{ij} + A_{w\,ji} + e_{ij}$$

where μ represents the overall mean, *ApY* is the fixed effect of the combination of the ith apiary (*i*= 1, 2, 3) and j^{th} year (*j*=1, 2); A_w represents the random genetic effect the j^{th} colony where the number of colonies per apiary varies between 17, 39, and 95, 151 in total (*j* = 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:

$$y_{ijkl} = \mu + ApTY_{ijk} + A_{w\,lkij} + pe_l + e_{ijkm}$$

where μ represents the overall mean, *ApTY* is the fixed effect term representing the combination of the i^{th} apiary (*i*= 1, 2, 3), j^{th} recording time (*j*= 1, 2) and k^{th} year of observation (*k*= 1, 2); A_w represents the random genetic effect the l^{th} colony (151 colonies in total), *pe* represents the random permanent environmental effect 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

Table I 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 and was safer, 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. 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 (Fig. 1 C and D, blue circles), giving complete control over the amount of killed brood.

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.

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.

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 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.

The phenotypic correlation in Figure 2 shows the similarity of 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 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 effect 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 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 (Fig. 1 C).

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

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

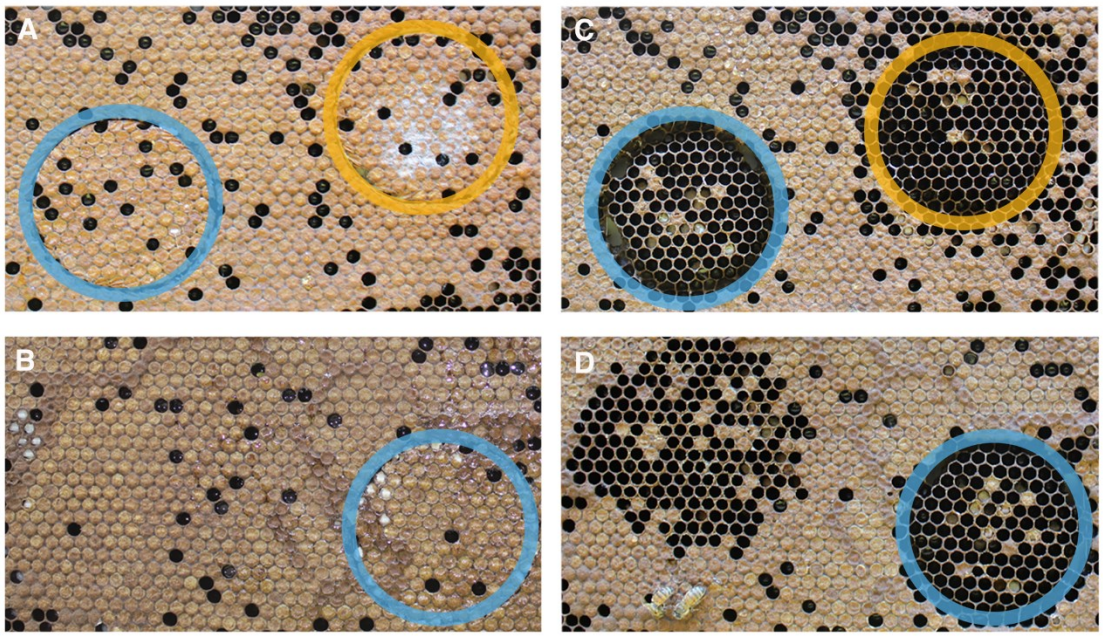

Figure 1 Visual comparison of the two methods. The pictures show the two sides (A and B) of a tested comb. Blue circles indicate FKB, orange circles indicate FKB. A) side A of the comb time zero, B) side B of the comb at time zero, C) side A of the comb after 24hours from the test, D) side B of the comb after 24 hours from the test.*

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 (rP), correlation of colony effects (rC) and correlation of error term (re) 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)	

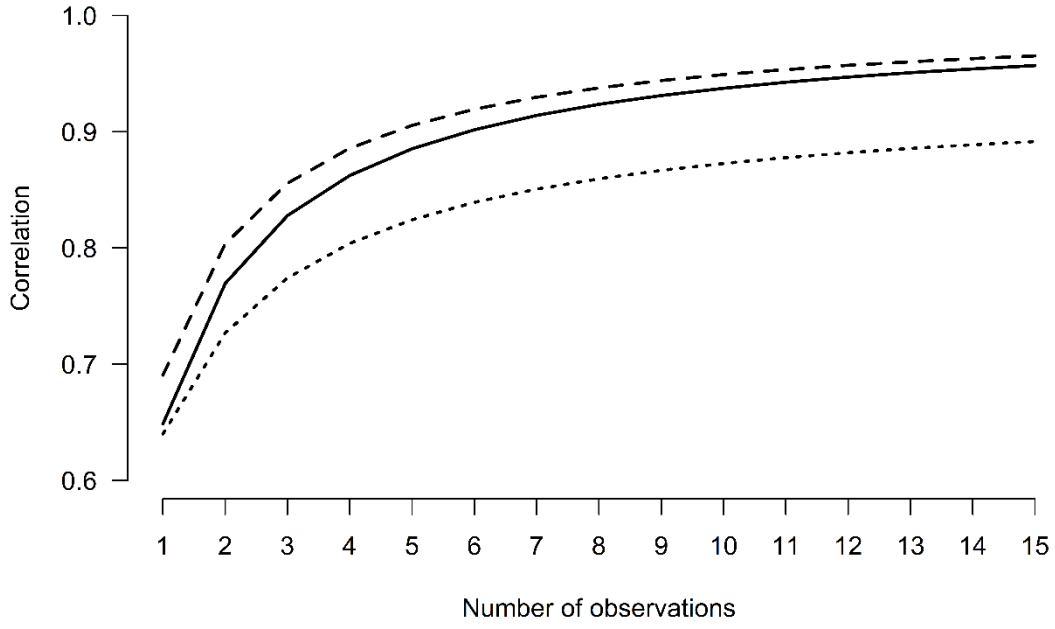

Figure 2 Correlations for mean of repeated HB records on a colony as a function of the number of observations. Dotted line: phenotypic correlation between FKB and FKB calculated with parameters from bivariate model; solid line: accuracy for FKB calculated with parameters from the univariate model; dashed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

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 ± 0.25) was higher than the one estimated with the repeatability model (0.23 ± 0.16), but it is more or less half of the only value (0.65 ± 0.61) reported in the literature for freeze killing recording method (Harbo and Harris, 1999). 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 very small ($2.7*10^{-4}$). 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 very similar suggesting that there is in principle no benefit of a repeatability model over an average model.

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*). $\bar{r}_{\hat{A},A}$ 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

4. Conclusion

FKB and FKB* essentially measure the same trait. 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 EBVs, there was no benefit of using a repeatability model over the use of a model for the average of two HB score.

Acknowledgements

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.

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 figures and tables. PB, RR and EWB provided critical feedback on the manuscript and participated in the revision of it.

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.

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

Applying the definition of correlation, it follows that:

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

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

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

Division of numerator and denominator by σ_C gives

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

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. DOI: *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

Butler, D. (2009) Asreml: asreml() fits the linear mixed model. R package version 3.0. *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 Montaño, 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

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

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(5), 792-796

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/

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, 405–42