



Genetic parameters, trends, and inbreeding in a honeybee breeding program for royal jelly production and behavioral traits

Benjamin BASSO^{1,2} , Tristan KISTLER^{2,3,4} , and Florence PHOCAS^{2,3} 

¹ INRAE, UR 406 Abeilles et Environnement, 84914 Avignon, France

² UMT PrADE, 84914 Avignon, France

³ Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France

⁴ Wageningen University & Research, Animal Breeding and Genomics, P.O. Box 338, 6700 AH Wageningen, The Netherlands

Received 6 September 2023 – Revised 12 December 2023 – Accepted 21 December 2023

Abstract – An economically sustainable production of royal jelly (RJ) requires the use of genetics specifically improved for this very demanding trait. The French Royal Jelly Association has developed since 2010 a selection plan for that purpose, based on a collective data recording system and the initial seedstock of five French beekeepers. The aim of this study was to evaluate the efficiency of this collective breeding program, a decade after its concrete start. We considered three different genetic models to describe the phenotypes, hereafter named the Colony, Worker, and Queen Models. We showed that when the size and the structure of the dataset are insufficient to fit a Colony Model, fitting a Queen Model gives more understandable results for practical use of the estimated breeding values than the Worker Model. Regarding genetic parameters, we estimated moderate to high heritability values for all traits, regardless of the genetic model. Under a Queen Model, heritability estimates of RJ production were 0.19 ± 0.09 and 0.35 ± 0.13 for first and second harvests, respectively; the estimates were 0.76 ± 0.08 , 0.29 ± 0.09 , and 0.22 ± 0.08 , respectively, for the calmness, gentleness, and spring development of the colony. We also proved that the breeding program is efficient, increasing the production at second harvest of 3.5% per year on average. Attention must, however, be paid to the evolution of inbreeding in this very small breeding population. Few genetic correlations were significantly different from zero between the traits, due to large standard errors of our estimates. It is, however, worth to note that the trends were generally favorable between all traits, which appears encouraging to improve production and behavioral traits jointly in the breeding program.

Apis mellifera / heritability / genetic correlation / genetic gain / breeding value / selective breeding / royal jelly

1. INTRODUCTION

Royal jelly (RJ) production is a very particular as it requires specific conditions (Li 2000). A full-time person can hardly manage more than 40 colonies as it is a highly time-consuming work per colony, but with a short 3-month harvesting season, generally from April to June in France. In

consequence, beekeepers need strains of honeybees highly improved for RJ production to reach a sufficient productivity per colony and to get a sufficient income from their activity. To our knowledge, apart from our preliminary estimates of genetic parameters for the breeding program of the French Royal Jelly Association, hereafter named GPGR (Phocas et al. 2017; Basso et al. 2022), little information is available in the technical or scientific literature on bee breeding programs for RJ production and comes mainly from China (Li

Corresponding author: B. Basso, benjamin.basso@inrae.fr
Manuscript editor: David Tarpay

2000; Chen et al. 2002; Li and Aiping 2005) or Brazil (Ostrovski-Tomporoski et al. 2016). All these authors reported considerable improvement of RJ production through selective breeding.

To better promote the value of their product in France, five producers created the GPGR Association in the mid-90s. Since then, they have strongly selected their colonies for this production until they managed a very specific population, clearly differentiated from others breeds in France as revealed by genomic analysis (Wragg et al. 2016). The initial breeding plan was based on mass selection within beekeepers' own seed-stock. Since 2011 it has evolved through collective data recording, queen testing, and within-family selection to preserve the maternal lines' genetic variability (Kistler et al. 2021).

For over 30 years in most farmed species, selection is no longer limited to mass selection, but integrates phenotypic information from all related animals to the selection candidates thanks to genetic and statistical models allowing the best linear unbiased prediction (BLUP) of breeding values (Henderson 1975). Due to the biological peculiarities of *Apis mellifera* (haplodiploidy, queen polyandry, and colony social phenotypes), the standard BLUP genetic evaluation models for livestock species could not be applied for honeybee evaluation. However, some BLUP methodologies and genetic models have been developed and used for the genetic evaluation of honeybee queens (Bienefeld et al. 2007; Brascamp and Bijma 2014, 2019a, b; Andonov et al. 2019; Guichard et al. 2020). Since 2017, RJ producers are using BLUP-estimated breeding values to rank their selection candidates considering the performance of the colony as a trait of the queen (Phocas et al. 2017), based on a so-called Queen Model in Guichard et al. (2020). Other genetic models used are either a Worker Model (Guichard et al. 2020; Du et al. 2022) or what was recently named a Colony Model (Du et al. 2022) that includes both the Queen and the Worker genetic effects in the BLUP genetic model as initially proposed by Bienefeld et al. (2007).

In order to ascertain the efficiency of the GPGR breeding program, the aims of the study were triple while accounting for the three

different genetic models to describe colony performances: (i) to estimate genetic parameters of RJ production and behavioral traits; (ii) to estimate realized genetic gains; and (iii) to assess the evolution of inbreeding in the population over the last decade.

2. MATERIAL AND METHODS

2.1. Honeybee population

Despite the fact that the GPGR data collection started in 2011, only phenotype records from queens born in years 2012 to 2021 were considered in the study (Supplementary Table S1), as consistent performance recording was only implemented since 2013. Our dataset includes information from 1420 queens of which 799 were open-mated and 621 were instrumentally inseminated.

In the GPGR breeding program, a new generation of 80 to 180 queens is produced each year from 6 to 23 1-year old breeding queens (BQs) (Supplementary Table S1). Some of them (29 to 109 virgin queens depending on the year) are instrumentally inseminated to renew the BQ pool, leading to a 1-year generation interval on the dam path. Other queens are naturally mated and part of them are used as drone-producing queens (DPQs) the year after testing. The generation interval is therefore of 2 years on the sire path. From 2020, natural mating queens are still produced but no longer phenotyped (Supplementary Table S1). Since then, the breeding program is only based on BQ phenotyping.

The same six experienced bee breeders working together since 2010 are responsible for the BQ selection nucleus and renew each year the breeding stock by exchanging genetic origins to connect their selection apiaries. In part due to this organization, only one DPQ is used by each bee breeder to inseminate the selection candidates and, when the same paternal line (i.e., same dam of the drone-producing queens) is used by different bee breeders, different DPQs are involved. On the other side, a single bee breeder produced each year all the queens which

are naturally mated at the same period and place and then tested by about ten different beekeepers (8 to 18 depending on the year).

2.2. Performance testing

To ensure that all workers of each colony are daughters of the tested queens, testing starts for all surviving queens after the first winter losses (Büchler et al. 2013), whether they are potential BQs or potential DPQs. The BQs are phenotyped for all traits in their birth apiary. For DPQs, the evaluation program starts each year with production of groups of sister queens (on average 11 queens per paternal line). These groups of potential sires are sent for evaluation to 8 to 18 test apiaries each year. Queens are randomly assigned to the various beekeepers so that each paternal line is distributed across three or four apiaries and each apiary is testing three different lines with three sisters per line.

Evaluation of the queens starts in spring of the following year, using all the data records collected for BQs by bee breeders and for tested DPQs by beekeepers. As early as possible in the season, between April and May, depending on the year and the area, beekeepers start the RJ production. For this, each hive has a queen-less part in which queen cells with young larvae (stage L1) are introduced. Then, every 3 days, after removal of the queen larva, beekeepers harvest the royal jelly deposited by worker bees, using aspirators specifically designed and manufactured for GPGR beekeepers. All the queen cells from a same hive are weighed together, with 1-g precision scale, before and after royal jelly extraction, to evaluate the quantity of jelly produced by each colony. New grafted queen cells are then immediately introduced. Between 30 and 120 cell cups are introduced at the same time, depending on the colony strength and on the proportion of queen cells it accepted to rear on the previous run. If the grafting is well done with very young larvae, this 3-day-interval corresponds to the optimum between the RJ quantity not consumed by larvae (that can be harvested) and the time spent to harvest it (Chen et al. 2002).

The recorded RJ production traits consist of the first two harvests in the production season, respectively P1 and P2. P1 is the production onset. P1 is therefore not necessarily representative of the capacity of the colony to produce over the rest of the season. In contrast, it has been shown (unpublished results) that the second harvest P2 corresponds more to the cruise production rate. Limiting the phenotyping to these two first harvests enables an early selection of queens during the production season.

During the production period, gentleness (GENT) and calmness during inspection (CALM) were evaluated with a visual score ranging from 1 (worst score) to 4 (best score) according to the Beebook scoring system (Büchler et al. 2013). Due to strong differences in the number of repeated records between colonies (the median being one with one to nine records per colony), we only considered the first set of scores given in a year for all colonies of an apiary. It corresponds to scores recorded within a single week for each apiary and production period. Spring development (DEVE) was evaluated once per year, at the end of March (before the start of the production period). DEVE aims to evaluate the strength of a colony through its brood surface in comparison of the strength of all the other colonies of the same apiary. Beekeepers scored 1 the weakest colony and 4 their strongest ones.

2.3. Pedigree preparation

In the GPGR pedigree file, each record consisted of a unique identification code (ID) for the colony's queen, the queen birth year, the queen breeder, the ID of the queen's dam, and the testing apiary in case of a tested colony. If known, the identification of the paternal line used to mate the queen was also included in the pedigree file through the ID of the dam of the DPQ that produced the drones used to inseminate the queen.

In order to get all the information required to build the relationship matrix for BLUP evaluation (Brascamp et Bijma 2014), we considered that the number of DPQs that inseminated a queen (NS

parameter) was equal to 1 for any potential breeding queen and the number of drones (ND parameter) mated to the queen was equal to 10. Unfortunately, the ID of DPQ used for insemination was not recorded, even if the DPQ was performance tested. A virtual ID of the corresponding DPQ was created through the combination of its dam ID and breeder's name since each breeder used a different DPQ from any given dam (corresponding to a given paternal line). For any uncontrolled mated queen (i.e., a potential DPQ), the number of drone mates (ND) was set to 20 and they were supposed to come from a pool of NS=100 DPQs assumed to come from unknown parents to mimic natural mating conditions (Tarpy and Nielsen 2002; Schlüns et al. 2005; Delaney et al. 2011; Tarpy et al. 2013).

The complete GPGR pedigree data included 4189 queens. Dam pedigree of all queens was traced back as far as possible. The founder population was composed of 22 queens born in year 2009, and there were 107 BQ with an average family size of 11 daughter queens (ranging from 1 to 37).

The pedigree file used for the BLUP genetic evaluation was created in two steps in order to run an animal model with an additional maternal genetic effect on the colony performance, i.e., to jointly evaluate the effects of the group of worker bees and of the colony's queen as proposed by Brascamp and Bijma (2014, 2019b). Firstly, the file containing records for a unique identification number generated for the group of workers within each colony was created followed by a sire ID, dam ID (queen ID of the colony), and test year of that particular colony. Secondly, it contained raw records from the queen ID, its own sire ID, its own dam ID, and queen birth year to complete all ancestral entries. The complete BLUP pedigree file included 10,868 records.

The inverse of the pedigree relationship matrix between all entries in the pedigree was calculated using an R program (R Core team 2017) developed by Brascamp and Bijma (2019a).

2.4. Statistical analysis

Genetic parameters and genetic trends were estimated considering three different BLUP animal

models to account for the genetic components playing a role on colony performance.

The model which is supposed to most accurately represent the honeybee biology is called the Colony Model (CM) in accordance with Du et al. definition (2022). This model takes into account a direct genetic effect and a maternal genetic effect.

It assumes that the phenotype P_{ijk} of a colony composed of a worker group of bees j and a dam k is described by the following equation:

$$P_{ijk} = \mu + CG_i + d_j + m_k + e_{ijk} \text{ (CM)}$$

where μ is the general performance mean, CG_i is the fixed effect of the contemporary group i (i.e., for behavioral traits, the combination of recording year \times apiary effects and for the production traits the combination of recording year \times apiary \times harvesting period effects), d_j is the direct genetic effect of the average worker bee j , m_k is the maternal genetic effect of the dam k , and e_{ijk} is the CM's residual effect.

The CM is similar to the classical model used in all livestock species as an Animal Model accounting for maternal genetic effects (Bienefeld et al. 2007).

The difference is that, in the Colony Model, the animal does not represent a single individual but the average of a group of workers.

The two other models we used are simplified models that were used because of convergence and inconsistency issues on small datasets (Guichard et al. 2020). They are used here for the very same reasons, as the restricted maximum likelihood algorithm did not converge to estimate genetic parameters under CM for behavioral traits, and CM estimates for production traits may deserve cautious attention (see Section 4). Under these models, the colony performance is only attributed to a single genetic effect, either the one of the worker group j or to the one of its queen k . This genetic effect has to be seen as a composite value that will capture genetic effects expressed by both the worker group and the queen without disentangling them (Du et al. 2022). In accordance with Guichard et al. (2020) and Du et al. (2022), we name these two models Worker Model (WM) and Queen Model (QM), respectively.

Under the Queen Model,

$$P_{ijk} = \mu + CG_i + q_k + \varepsilon_{ijk} \text{ (QM)}$$

where q_k is the queen effect and ε_{ijk} is the residual effect in the Queen Model.

According to Du et al. (2022) formula, we can derive $q_k = m_k + 0.5 d_k$.

Under the Worker Model,

$$P_{ijk} = \mu + CG_i + w_j + \gamma_{ijk} \text{ (WM)}$$

where w_j is the average genetic effect of worker group j and γ_{ijk} is the WM's residual effect.

According to Du et al. (2022) formula, we can derive

$$w_j = d_j + (0.5/a_{ww}) m_j$$

where a_{ww} corresponds to the average relationship coefficient between two workers in a colony (Brascamp and Bijma 2014, 2019a) and was estimated equal to 0.3915 in our average population and equal to 0.2775 in our base population (see Appendix A for calculation).

The phenotypic variance σ_p^2 emerges from both the maternal genetic variance due to the single queen and the variance of the average genetic effect of the workers of the colony. Therefore, the genetic variance attributed to the worker group in the expression of σ_p^2 is reduced by the coefficient a_{ww} compared to the variance attributed to a single worker (Brascamp and Bijma 2018, 2019a).

Therefore, in the Colony Model, the phenotypic variance is partitioned as

$$\sigma_p^2 = \sigma_m^2 + a_{ww}\sigma_d^2 + \text{Cov}_{dm} + \sigma_e^2$$

with σ_m^2 the maternal genetic variance, σ_d^2 the direct genetic variance, Cov_{dm} the covariance between direct and maternal genetic effects, and σ_e^2 the CM's residual variance.

Under CM, two heritability estimates were derived—one for the direct genetic effect of the workers ($h_d^2 = \sigma_d^2/\sigma_p^2$), and another for the maternal genetic effect ($h_m^2 = \sigma_m^2/\sigma_p^2$)—while the sole heritability of the trait attributed to the queen effect (h_q^2) was estimated under QM, and the sole heritability of the trait attributed to the worker genetic effect (h_w^2) was estimated under WM.

To compare the consistency of estimates of genetic parameters between the three models, CM, WM, and QM, we recalculated the queen and worker genetic variances using CM estimates following equations detailed in Appendix B.

Bi-trait QM and WM analyses were also run to estimate genetic correlations between all traits. All analyses were performed using the AIREMLF90 program from the BLUPf90 package (Misztal et al. 2002).

3. RESULTS

3.1. Performance testing results

The potential BQs, which are inseminated queens, show a higher RJ production on average than the potential DPQs, that are naturally mated (Table I). The average increase was +30% for P1 and even more for P2 (+36%). On the contrary, coefficients of variation of production traits were higher (~55%) for DPQs than those of BQs (~45%). The average grades on behavior traits were more similar between BQs and DPQs, but with a 20% higher variability of the scores for GENT and CALM for DPQs than for those of BQs.

Table II shows the mean and standard deviation of raw performances, as well as the phenotypic variances estimated after correction of raw performances for the fixed contemporary group effects and the phenotypic correlations between these corrected performances.

Ratios of phenotypic to raw variances indicate that the fixed effects (i.e., recording year \times apiary effects) explained 50 to 53% of the variance of raw performances for all traits, except 32% for DEVE.

3.2. Estimates of genetic parameters for production traits

Table III presents the estimates of variance components for the three BLUP models CM, QM, and WM.

Regarding estimates of phenotypic variances derived as the sum of genetic and residual variances (see equations 1 to 3 in Appendix B), they

Table I Performance of breeding queens (BQ) and drone-producing queens (DPQ) and of full data set for all traits (P1: production at first harvest, P2: production at second harvest, GENT: gentleness, CALM: calmness during inspection, DEVE: spring development)

		BQ			DPQ			Full data set		
		<i>N</i>	Mean	SD	<i>N</i>	Mean	SD	<i>N</i>	Mean	SD
Production	P1 (g)	598	37.11	17.00	784	28.50	15.76	1382	32.21	16.84
	P2 (g)	596	42.09	18.47	767	30.89	16.99	1363	35.79	18.50
Behavior	GENT	546	2.99	0.56	639	2.94	0.68	1185	2.96	0.63
	CALM	547	3.09	0.51	637	3.08	0.66	1184	3.09	0.60
	DEVE	437	2.93	0.89	603	3.00	0.87	1040	2.97	0.88

varied strongly across the different genetic models (Table III) and were also markedly different from the variances derived for P1 and P2 only after correction of raw performances for the fixed contemporary group effects (Table II).

Phenotypic variances estimated between the CM and WM were very close. However, the estimates under QM were clearly lower and closer to the variances derived ignoring any genetic effects (Table II).

In addition, the phenotypic variance estimated under a WM depends on the coefficient a_{ww} . Assuming the coefficient derived for naturally mated queens (i.e., $a_{ww} = 0.2775$ as it happens in the base population), estimates were 155.22 and 195.79 for P1 and P2, respectively (Table III), while assuming the average a_{ww} for all the phenotyped colonies in the recorded population; the

corresponding estimates were 164.06 and 209.65 (Supplementary Table S2).

Assuming $a_{ww} = 0.2775$, the heritability estimates under CM were moderate for P1 (respectively 0.26 and 0.44 for the maternal genetic effects of the queen and the direct genetic effects of the workers) and high for P2 (respectively 0.53 and 0.78), but always higher for the worker effects than the queen effects. For both traits, the direct-maternal genetic correlations estimated under CM were strongly negative but with very large standard errors (Table III), rendering them neither significantly different from 0 or -1. Assuming $a_{ww} = 0.3915$ had little impact on heritability estimates under CM (Supplementary Table S2).

Under QM, heritability estimates were 0.19 and 0.35, respectively, for P1 and P2. Those values were lower than the heritability of maternal genetic effects under CM. Under WM, the heritability for P1 was estimated at a 5% higher value than the heritability of direct genetic effect under CM. On the contrary, the heritability for P2 under WM was estimated 20% lower than the heritability of direct genetic effects under CM.

Table II Variances on the diagonal and correlations above the diagonal for performances corrected for the fixed contemporary group effects estimated by ANOVA model (P1: production at first harvest, P2: production at second harvest, GENT: gentleness, CALM: calmness during inspection, DEVE: spring development)

	P1	P2	GENT	CALM	DEVE
P1 (g)	135.79	0.63	0.03	0.05	0.15
P2 (g)		171.20	0.06	0.09	0.11
GENT			0.19	0.39	-0.03
CALM				0.17	0.01
DEVE					0.52

3.3. Breeding value estimation for production traits

Figure 1 shows the annual trends for P2 performances as well as maternal and direct EBV standardized by genetic standard deviation for BQ and DPQ under CM.

Table III Estimates of variance and heritability values (h^2) under respectively colony model (CM), queen model (QM), and worker model (WM) for production at first harvest (P1) and second harvest (P2) (\pm SE), considering only maternal half-sisters as workers in the colonies ($a_{ww}=0.2775$)

	P1			P2		
	CM	QM	WM	CM	QM	WM
Phenotypic var	155.63 \pm 8.16	161.30 \pm 8.25	155.22 \pm 12.67	201.76 \pm 12.37	212.76 \pm 12.29	195.79 \pm 8.61
Residual var	130.91 \pm 8.77	134.35 \pm 8.96	133.72 \pm 8.28	145.33 \pm 12.60	153.47 \pm 12.35	162.04 \pm 10.32
Genetic var. ^a	68.60 \pm 40.62	26.95 \pm 12.32	77.48 \pm 31.61	156.61 \pm 70.82	59.29 \pm 19.63	122.62 \pm 41.15
Maternal genetic var	40.62 \pm 26.88	Na	Na	107.59 \pm 51.70	Na	Na
h^2	0.44 \pm 0.27	0.18 \pm 0.08	0.50 \pm 0.20	0.78 \pm 0.36	0.32 \pm 0.10	0.62 \pm 0.20
h^2_m	0.26 \pm 0.17	Na	Na	0.53 \pm 0.25	Na	Na
r_{dm}	-0.66 \pm 0.91	Na	Na	-0.73 \pm 0.39	Na	Na

^aThe genetic variance is the direct genetic variance due to the worker bee group under CM, the queen genetic variance under QM, and the worker genetic variance under WM. For the CM, h^2_m and r_{dm} denote respectively the heritability for maternal effects and the genetic correlation between direct and maternal genetic effects

Yearly variations in production were high, but the average performance for P2 increased respectively by 0.74 g and 1 g per year for BQ and DPQ during the observed period from year 2012 to year 2021.

Regarding EBVs based on CM, estimated direct genetic gains were two times higher for the inseminated BQ than for DPQ whose paternal path is uncontrolled (Figure 1). For BQ, direct genetic trend was estimated to 0.60 g/year, corresponding to an annual increase of $\sim 7\%$ of direct genetic standard deviation from year 2012 to year 2021. Meanwhile, no genetic trends were observed for maternal breeding values that stay stable over years for both BQ and DPQ direct breeding values (Figure 1). Note that maternal EBV reliabilities were significantly higher than direct EBV ones (Supplementary Table S3).

When considering EBVs estimated under QM and WM for P2 (Figure 2), similar differences in genetic trends were observed between DPQ and BQ as under CM with an annual increase for BQ EBVs of 8 and 9% of genetic standard deviation, respectively, under QM and WM.

3.4. Phenotypic variance and heritability estimates for behavioral traits

Regarding estimates of phenotypic variances for behavioral traits, they show the same trend as

production traits, with strong differences across the two genetic models (Table IV) and are also markedly different from the variances derived for raw performances only corrected for the fixed contemporary group effects (Table II).

Heritability estimates under QM were at moderate values for DEVE and GENT, but at a very high value of 0.76 for CALM. As for production traits, WM estimates were larger with high heritability values for the three traits. All WM estimates were reduced by 10% when considering $a_{ww} = 0.3915$ (Supplementary Table S4).

3.5. Phenotypic and genetic correlations across all traits

The phenotypic correlation between P1 and P2 was estimated about 0.60–0.65 either on raw performances corrected for the contemporary group effects or under QM or WM BLUP genetic models (Table II; Supplementary Tables S5 and S6). Those two traits were genetically highly correlated, with estimates of correlation ranging from 0.72 under WM (Supplementary Table S6) to 0.90 under QM (Supplementary Table S5).

Phenotypic correlation between production traits and behavioral traits were close to zero regardless of the estimation method (Table II; Supplementary Tables S5 and S6). Due to large standard errors of the estimates,

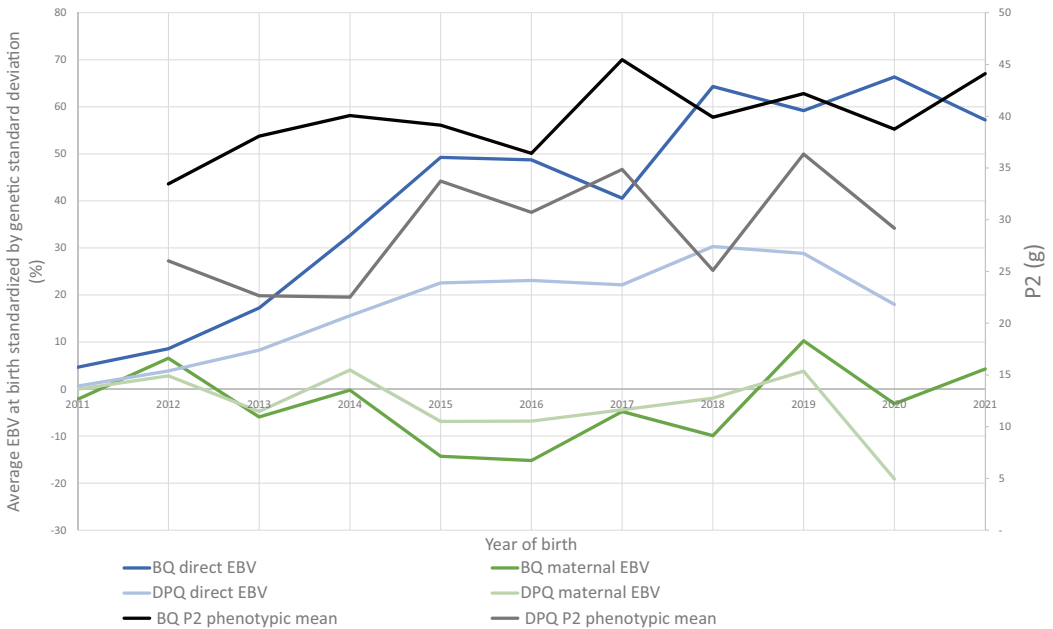


Figure 1. Average phenotypic values, average direct and maternal estimated breeding values (EBV) for breeding queens (BQ) and drone-producing queens (DPQ) under the colony model for production at second harvest (P2) depending on queen birth year.

very few genetic correlations were significantly different from zero between behavioral traits or between these former and production traits. The only exceptions were the significant positive genetic correlations between P2 and CALM (under QM) or GENT (under QM and WM) as well as the one between CALM and GENT under QM (0.94) (Supplementary Table S5) and under WM (1.00) (Supplementary Table S6). Those two behavioral traits had very different estimates of genetic correlations with the two production traits despite the high genetic correlation estimated between P1 and P2; while positive genetic correlations were estimated with P2, negative estimates, not significantly different from 0, were derived with P1.

3.6. Estimated genetic trends for all traits and inbreeding

Regardless of the models, genetic improvement for all traits was observed (see Supplemental Figure S1), except for DEVE under WM. In the last 10 years, considering WM evaluation, P2 and CALM were

improved by almost one genetic standard deviation while P1 and GENT were improved by 0.6 genetic standard deviation. Considering QM evaluation (that gave a greater weight to maternal genetic effects than to the direct genetic effects of the workers), all traits were improved in between 0.30 and 0.90 genetic standard deviation during these last 10 years.

In the same time, inbreeding increased in average by 0.8% each year from 2015 to 2021 (result not shown). As inbreeding is only derived based on the available pedigree, the inbreeding rate could not be correctly estimated for the first 4 years of the breeding program due to insufficient pedigree depth.

4. DISCUSSION

4.1. Efficiency of the GPGR breeding program

Since 2012, we observed a clear improvement of RJ production (based on P2 records), from

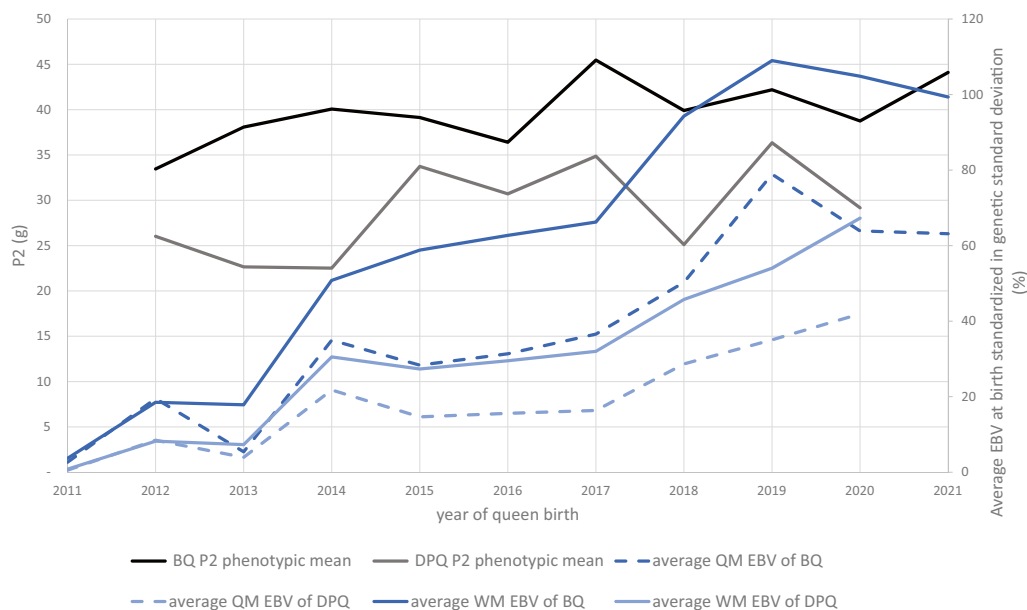


Figure 2. Average phenotypic values and corresponding average estimated breeding values (EBV) for BQ and DPQ with queen model (QM) and worker model (WM) for production at second harvest (P2) depending on queen birth year.

about 34 to 46 g for breeding queens born in 2021 (Figure 1). However, the phenotypic gain seems far much less than the one reported for the Chinese breeding program that was initiated in the 1980s to derive a high royal jelly-producing honeybee strain from Italian bees (*A. mellifera ligustica*). At the start of the breeding program, production of RJ was about 20 g over a 3-day

period and increased to about 150 g per colony over 3 days in the early 2000s (Chen et al. 2002). In particular, a strong improvement from 30 g in 1988 to 100 g over 3 days in 1993 was observed due to an intensive selection applied over six generations at Zhejiang University to create the ‘Zhenongda No. 1’ high RJ line from a founder population of 17 lines from 1000 high-yielding

Table IV Estimates of variance components and heritability (h^2) and associated standard errors under queen model (QM) and worker model (WM) for behavior traits (GENT: gentleness, CALM: calmness during inspection, DEVE: spring development), considering only maternal half-sisters as workers in the colonies ($a_{ww}=0.2775$)

	GENT		CALM		DEVE	
	QM	WM	QM	WM	QM	WM
Phenotypic var	0.24 ± 0.02	0.22 ± 0.01	0.30 ± 0.03	0.21 ± 0.01	0.61 ± 0.03	0.59 ± 0.03
Residual var	0.17 ± 0.015	0.18 ± 0.01	0.07 ± 0.02	0.14 ± 0.01	0.48 ± 0.04	0.49 ± 0.04
Genetic var. ^a	0.07 ± 0.03	0.13 ± 0.05	0.23 ± 0.04	0.26 ± 0.08	0.13 ± 0.05	0.35 ± 0.15
h^2	0.29 ± 0.09	0.59 ± 0.21	0.76 ± 0.08	1.23 ± 0.32	0.22 ± 0.08	0.60 ± 0.23

^aThe genetic variance is the queen genetic variance under QM and the worker genetic variance under WM

colonies recorded across various high RJ-producing regions of China. We may hypothesize that the major part of this reported increase was jointly due to the extreme selection pressure to choose the 17 lines as well as a strong technical improvement in managing RJ colonies (Chen et al. 2002). In particular, French beekeepers have to respect the ISO12824:2016 standard (AFNOR 2016) for RJ production which forbid to give sugar during production. Such input can highly increase production but cause change in RJ composition (Wytrychowski et al. 2013).

As far as we know, our study is the first one trying to disentangle genetic and environmental factors that may explain the observed phenotypic increase in RJ production through the estimation of genetic trends. Despite the fact that our estimates strongly varied depending on the genetic model we considered to predict breeding values, all estimates consistently showed a strong increase in direct genetic effects of workers while no clear improvement of maternal genetic effects could be observed neither under CM (Figure 1) or comparing QM EBVs with WM EBVs (Figure 2).

Since 2017, a BLUP-QM evaluation is routinely performed to sort selection candidates of the GPGR breeding population. However, the final choice is still mainly taken through collective discussion of the bee breeders, using mainly phenotypic information for RJ production and choice within maternal families. Comparing to simulations done by Kistler et al. (2021), the operational GPGR breeding program is very close to the simulated scenario “Within family mass selection & Polyandrous mating.” In this scenario and considering a higher heritability for maternal effects than for worker direct effects, Kistler et al. (2021) predicted a higher improvement for the worker direct genetic effects than for the maternal genetic effects, which is consistent with our EBV estimates. Indeed, the variance of direct effects explaining a larger part of the phenotypic variance, selecting on phenotypes induced a stronger improvement on direct effects than on maternal ones. Kistler et al. (2021) also showed that, when a negative genetic correlation existed between direct and maternal genetic effects, the difference between the corresponding genetic gains was exacerbated.

Despite the lesser importance given to behavioral traits in the selection choices made by GPGR, they showed positive estimated genetic trends for both QM and WM EBV. Under QM, the annual direct genetic gain for CALM was estimated to 52.3% of genetic standard deviations from year 2013 to year 2021 while evolution of the maternal genetic trend was 35.5% per year. The only exception being a null trend for DEVE breeding values estimated under WM. This exception must be explained by the very negative genetic correlations estimated under WM between DEVE and P2 which is the main selection criterion (Supplementary Table S6). As the corresponding estimates were significantly closer to zero under QM, this leads us to hypothesize that the direct genetic effects for DEVE and RJ production traits are negatively correlated while probably the maternal genetic effects are not (or at least in a lesser extent).

We also gave a pedigree-based estimate of inbreeding level in the population as it is a strong concern for a small breeding population managed as an almost closed nucleus for 10 generations of selection. Using pedigree rather than genomic information, we underestimated inbreeding levels, especially the first 4 years of the breeding program as the inbreeding coefficients were directly related to the amount of pedigree information which is available. However, then we could properly estimate an annual increase of inbreeding of 0.8% from year 2015 to year 2021, which corresponded to a rate of inbreeding increase of 1.2% per generation. This estimate was close to the annual estimate (1.1%) obtained by simulation of a similar breeding program (Kistler et al. 2021). Average inbreeding coefficients for BQ born in years 2019–2021 were ranging from 7.6 to 9.5%. Recently, Druml et al. (2023) estimated inbreeding levels within the two Austrian Carnica breeding populations selected through EBVs for over the last 30 years; inbreeding of queens with full pedigree recorded over five generations was estimated about 5.5% on average in the populations of active breeding queens for the ACA (2675 BQ in 2019) and ZAC (1286 BQ in 2020) breeding programs. Both populations exhibit very moderate loss of

genetic diversity in comparison to the GPGR one whose population size is very small with only 109 queens inseminated in 2020 (Supplementary Table S1).

The GPGR association should try to significantly increase the size of their breeding population to limit the risk of collapse. It is advisable to limit increasing inbreeding by 0.5 to 1% per generation in order to preserve long-term genetic gains as well as the adaptive capacities of the population in the event of a change in the environment (Felsenstein 1965). To reach the highest possible genetic gains while limiting inbreeding rate in honeybee breeding populations, further improvement of the breeding program should consider an optimum contribution selection strategy (Meuwissen 1997).

4.2. Heritability estimates according to genetic models for production traits

Regarding all models and traits studied, RJ production clearly showed moderate to high heritability, which confirms the results of Basso et al. (2022) on an earlier analysis of the GPGR breeding program. To our knowledge, there is only one other study (Faquinello et al. 2011)

presenting heritability for the RJ production. This first estimate was at a very low value of 0.06 in a Brazilian small population of 174 colonies from four generations derived from 20 naturally mated Africanized queens of *A. mellifera*. It was based on a mixed Animal Model in which the animal was considered as the worker but the paternal pedigree was ignored. This first estimate was surprisingly low as in general production traits exhibit moderate to high heritability values whatever the farmed species considered. If we compare our estimate for RJ production to recent estimates obtained for honey production we were in the same range of the literature when considering the same genetic model (Table V).

Our negative estimate of the genetic correlation between direct and maternal genetic effects for RJ production is close to the one estimated for honey yield in a large population of Austrian Carnica bees (Brascamp et al. 2016) while, on the other hand, an almost null correlation was derived in a naturally mated Italian bee population (Andonov et al. 2019). In these two studies, higher heritability values for direct worker effects than for maternal ones were estimated for honey yield (Table V).

Therefore, it seems that regardless of the production trait (honey or royal jelly), direct

Table V Heritability estimates ($h^2 \pm \text{SE}$) in recent literature for honey yield, gentleness and calmness under various genetic models: queen model (QM), worker model (WM), and colony model (CM)

Reference	<i>A. mellifera</i> sub species	Population size	Model	h^2	HONEY	GENT	CALM
Guichard et al. (2020)	<i>mellifera</i>	1066	QM	h_q^2	0.10 ± 0.06	0.32 ± 0.08	0.12 ± 0.06
			WM	h_w^2	0.06 ± 0.09	0.85 ± 0.21	0.39 ± 0.17
Guichard et al. (2020)	<i>carnica</i>	1007	QM	h_q^2	0.11 ± 0.06	0.02 ± 0.04	0.09 ± 0.05
			WM	h_w^2	0.27 ± 0.13	0.04 ± 0.08	0.06 ± 0.09
Andonov et al. (2019)	<i>ligustica</i>	3931	CM	h_m^2	0.12 ± 0.01	0.48 ± 0.12	-
			CM	h_d^2	0.33 ± 0.03	1.03 ± 0.08	-
			QM	h_q^2	0.26 ± 0.04	0.36 ± 0.01	-
Brascamp et al. (2018)	<i>carnica</i>	14,948	CM	h_m^2	0.70 ± 0.14	0.19 ± 0.08	0.10 ± 0.06
			CM	h_d^{2a}	1.34 ± 0.16	0.57 ± 0.08	0.61 ± 0.08
Tahmasbi et al. (2015)	<i>meda</i>	10,000	QM	h_q^2	0.22 ± 0.02	0.44 ± 0.03	-
Zakour et al. (2012)	<i>syriaca</i>	1088	QM	h_q^2	0.27 ± 0.32	0.08 ± 0.28	-

^aHere, we derived h_d^2 as a measure of heritability for the scope of selection, $h_d^2 = h^2/a_{ww}$ where h^2 is the value reported by Brascamp et al. (2018)

genetic effects of the workers have more impact on colony performance than the maternal genetic effects of the queen, and that both genetic effects are negatively correlated. However, those estimates of negative correlations must be taken with caution as strong biases may be induced due to uncertainties in the paternal pedigrees of the colonies (Kistler et al. 2023) and lack of data and sufficient depth of pedigree to disentangle direct and genetic effects as already reported by Andonov et al. (2019) and Guichard et al. (2020).

To assess the consistency of our results between CM and WM, we tried to rederive the genetic variances we should estimate under WM if the true direct and maternal genetic variances were the ones we estimated under CM (see Supplementary Table S7). This estimate depends on the value of the relationship coefficient a_{ww} we assume for workers in a colony. Considering $a_{ww}=0.2775$, the recalculated worker genetic variance based on the CM variance estimates was close to the direct estimate under WM for P1, but it was strongly overestimated (+36%) for P2. Considering $a_{ww}=0.3915$, recalculated worker genetic variances were strongly underestimated for both P1 and P2. To use relevant estimates of variance components under a WM evaluation, it is tricky to choose a proper value for a_{ww} as, in most breeding plans, both insemination and natural mating are used in varying proportions depending on year opportunities, and the numbers of DPQ and drones used to mate queens also significantly vary.

The coefficient a_{ww} is not involved in the derivation of the queen genetic variance based on the CM variance estimates. Indirect and direct estimates of the queen genetic variances were close (see Supplementary Table S7). Therefore, when data are insufficient to fit a CM, we may recommend to run a QM evaluation rather than a WM one. However, it should be noted that, doing so, bee breeders will put more emphasis on improvement on maternal genetic effects rather than on direct genetic effects on the colony performance when selecting their replacement queens on QM EBV rather than on WM EBV.

4.3. Heritability estimates for behavior traits

Our estimates of heritability for behavior traits are in the very large range of values given during the last decade in the literature for similar traits and genetic models applied to other bee populations (Table V).

As regards to GENT (also called defensive behavior), several REML estimates of heritability were given in the literature during the last decade (Table V). These previous estimates varied from very large values such as over 1 for the worker direct heritability under CM to values under QM ranging from intermediate to low values (<0.10).

Our estimate (h^2_w) of heritability for CALM under WM, which are outside the range of usual values (i.e., above 1, see Table IV) is a fine illustration of the issues we have to face to present understandable estimates of genetic parameters. This value relates to the heritable variance of the effect of a colony considered as a single individual which is a measure for the scope of selection. Indeed, Bijma (2011) established that the heritable variance in traits that depend on genes from multiple individuals is not limited to phenotypic variance and can exceed the value of 1, which is a fundamental difference from classical theory. A second kind of heritability ($h^2 = a_{ww}h^2_w$ as reported by Brascamp et al. 2018; Guichard et al. 2020) relates to the worker effect considered as a group of bees, which reflects the part of the phenotypic variance due to the genetic effect of the worker group and has therefore 1 as theoretical upper bound.

As far as DEVE (also called spring development) is concerned, little information is available to compare our estimates with the literature. However, our heritability estimated under QM ($h^2_q=0.22$) is close to the one derived for a Canadian honeybee population ($h^2_q=0.30$, Maucourt et al. 2020). These moderate values estimated for the development of the colony at spring time should allow to efficiently select on this trait, which is a matter of importance to get colonies as strong as possible early in the production season since the period of honey flow tends to be reduced by climate change.

Concerning our estimates of genetic correlations, very few are significant because we only had 107 families of breeding queens with small family sizes in our study. Large sample sizes (at least several hundreds of families) are required to accurately estimate genetic correlations as shown by some empirical studies (see for instance Koots and Gibson 1996). As advised by some geneticists 30 years ago (Cheverud 1988; Roff 1995), it seems advisable to use the phenotypic correlations (presented in Table II) that are probably better estimates of the true genetic correlations, than the ones we tried to derive under the QM or WM BLUP models.

5. CONCLUSION

As frequently reported in the literature on genetic parameters for honeybees, we faced many doubts on the relevance of our estimates of genetic parameters because of uncertainty on paternal pedigrees and joint use of open mated and inseminated queens. When data is still insufficient to implement a BLUP-Colony Model to jointly estimate maternal and direct genetic breeding values, we recommend to run a Queen Model rather than a Worker Model to better control and understand both estimates of breeding values and genetic parameters.

Despite these challenges, our study clearly shows that selection on royal jelly production is feasible and efficient even in a very small population. Future development of the breeding scheme should try to limit the inbreeding rate through an increase in the breeding population size.

SUPPLEMENTARY INFORMATION

The online version contains supplementary material available at <https://doi.org/10.1007/s13592-023-01055-3>.

ACKNOWLEDGEMENTS

We thank Charlotte Le Bihan, Tommy Gerez, and all beekeepers involved in the breeding program of the “Groupeement des Producteurs de Gelée Royale Française” (GPGR) for providing all these data and our great collaboration since years.

AUTHORS' CONTRIBUTIONS

BB and FP designed the study, performed the data analysis and drafted the first version of the manuscript. BB, TK and FP interpreted the results and commented on previous versions of the manuscript. All authors read and approved the final version of the manuscript.

DATA AVAILABILITY

The datasets generated or analyzed during the current study are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

Not applicable.

DECLARATIONS

Ethics approval Not required. The work involved data collected for farming purposes on an unregulated and very common insect species (*Apis mellifera*).

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly

from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

REFERENCES

- AFNOR NF ISO 12824 « Gelée royale – Spécifications, V 36–001, 26 novembre 2016, éditeur AFNOR
- Andonov S, Costa C, Uzunov A, Bergomi P, Lourenco D, Misztal I (2019) Modeling honey yield, defensive and swarming behaviors of Italian honey bees (*Apis mellifera ligustica*) using linear-threshold approaches. *BMC Genet* 20(1):78. <https://doi.org/10.1186/s12863-019-0776-2>. PMID:31638899;PMCID:PMC6805448
- Basso B, Kistler T, Gerez T, Phocas F (2022) Genetic analysis of royal jelly production and behaviour traits of honeybees. *Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP)* 2560. https://doi.org/10.3920/978-90-8686-940-4_619
- Bienefeld K, Ehrhardt K, Reinhardt F (2007) Genetic evaluation in the honey bee considering queen and worker effects — a BLUP-Animal Model approach. *Apidologie* 38:77–85. <https://doi.org/10.1051/apido:2006050>
- Bijma P (2011) A general definition of the heritable variation that determines the potential of a population to respond to selection. *Genetics* 189(4):1347–1359. <https://doi.org/10.1534/genetics.111.130617>
- Brascamp EW, Bijma P (2014) Methods to estimate breeding values in honey bees. *Genet Sel Evol* 46(1):1–15. <https://doi.org/10.1186/s12711-014-0053-9>
- Brascamp EW, Bijma P (2019a) A note on genetic parameters and accuracy of estimated breeding values in honey bees. *Genet Sel Evol* 51(1):71. <https://doi.org/10.1186/s12711-019-0510-6>
- Brascamp EW, Bijma P (2019b) Software to facilitate estimation of genetic parameters and breeding values for honey bees. In *Proceedings of the 46th Apimondia 8–12 September 2019; Montréal, Canada*. <https://edepot.wur.nl/546515>. Last access 01 Sept 2023
- Brascamp EW, Willam A, Boigenzahn C, Bijma P, Veerkamp RF (2016) Heritabilities and genetic correlations for honey yield, gentleness, calmness and swarming behaviour in Austrian honey bees. *Apidologie* 47:739–748. <https://doi.org/10.1007/s13592-016-0427-9>
- Brascamp EW, Willam A, Boigenzahn C, Bijma P, Veerkamp RF (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, Andonov S, Bienefeld K, Costa C, Hatjina F, Kezić N, Kryger P, Spivak M, Uzunov A, Wilde J (2013) Standard methods for rearing and selection of *Apis mellifera* queens. In V Dietemann; J D Ellis; P Neumann (Eds) *The COLOSS BEEBOOK, Volume I: standard methods for *Apis mellifera* research*. *J Apicul Res* 52:1–29
- Chen S, Su S, Lin X (2002) An introduction to high-yielding royal jelly production methods in China. *Bee World* 83:69–77. <https://doi.org/10.1080/0005772X.2002.11099543>
- Cheverud M (1988) A comparison of genetic and phenotypic correlations. *Evolution* 42:958–968
- Delaney DA, Keller JJ, Caren JR, Tarpay DR (2011) The physical, insemination, and reproductive quality of honey bee queens (*Apis mellifera* L.). *Apidologie* 42:1–13
- Druml T, Putz A, Rubinigg M, Kärcher MH, Neubauer K, Boigenzahn C (2023) Founder gene pool composition and genealogical structure in two populations of Austrian Carniolan honey bees (*Apis mellifera carnica*) as derived from pedigree analysis. *Apidologie* 54:24. <https://doi.org/10.1007/s13592-023-00999-w>
- Du M, Bernstein R, Hoppe A, Bienefeld K (2022) Influence of model selection and data structure on the estimation of genetic parameters in honeybee populations. *G3 (Bethesda)* 4:12(2):jkab450. <https://doi.org/10.1093/g3journal/jkab450>
- Faquinello P, Arnaut de Toledo V, Nunes M, Oliveira C, Sereia M, Costa Maia F, Ruvolo-Takasusuki MC (2011) Parameters for royal jelly production in Africanized honeybees. *Sociobiology* 57:495–509
- Felsenstein J (1965) The effect of linkage on directional selection. *Genetics* 52(2):349–363. <https://doi.org/10.1093/genetics/52.2.349>
- Guichard M, Neuditschko M, Soland G, Grandjean FP, M, Gerster S, Dainat B, Bijma P, Brascamp EW, (2020) Estimates of genetic parameters for production, behaviour, and health traits in two Swiss honey bee populations. *Apidologie* 51:876–891. <https://doi.org/10.1007/s13592-020-00768-z>
- Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31:423–447. <https://doi.org/10.2307/2529430>
- Kistler T, Basso B, Phocas F (2021) A simulation study of a honeybee breeding scheme accounting for polyandry, direct and maternal effects on colony performance. *Genet Sel Evol* 53:71. <https://doi.org/10.1186/s12711-021-00665-8>
- Kistler T, Brascamp EW, Basso B, Bijma P, Phocas F (2023) Uncertainty in the mating strategy causes bias and inaccuracy in estimates of genetic parameters in honeybees. *bioRxiv* 541688. <https://doi.org/10.1101/2023.05.22.541688>
- Koots KR, Gibson JP (1996) Realized sampling variances of estimates of genetic parameters and the difference between genetic and phenotypic correlations. *genetics*. 143;3(1):1409–1416. <https://doi.org/10.1093/genetics/143.3.1409>
- Li J (2000) Technology for royal jelly production. *American Bee Journal* 140:469–472
- Li J, Aiping W (2005) Comprehensive technology for maximizing royal jelly production. *Am Bee J* 145:661–664
- Maucourt S, Fortin F, Robert C, Giovenazzo P (2020) Genetic parameters of honey bee colonies traits

- in a Canadian selection program. *InSects* 11:587. <https://doi.org/10.3390/insects11090587>
- Meuwissen THE (1997) Maximizing the response of selection with a predefined rate of inbreeding. *J Anim Sci* 75(4):934–940. <https://doi.org/10.2527/1997.754934x>
- Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH (2002) BLUPF90 and related programs (BGF90). In Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France 19–23
- Ostrovski-Tomporoski K, Faquinello P, Costa-Maia FM, Ruvolo-Takasusuki MC, Santos PR, Toledo VA (2016) Breeding program design principles for royal jelly. *INTECH*. <https://doi.org/10.5772/62417>
- Phocas F, Bihan CL, Basso B (2017) Genetic parameters of royal jelly production and behavioural traits of honey bees. In: Proceedings of the 68th annual meeting of the European Association for Animal Production (EAAP) Tallinn
- R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
- Roff DA (1995) The estimation of genetic correlations from phenotypic correlations: a test of Cheverud's conjecture. *Heredity* 74:481–490
- Schlüns H, Moritz RFA, Neumann P, Kryger P, Koeniger G (2005) Multiple nuptial flights, sperm transfer and the evolution of extreme polyandry in honeybee queens. *Anim Behav* 70:125–131
- Tahmasbi G, Kamali MA, Ebadi R (2015) Nejati javaremi A, Babaei M, Gharadaghi AA, Bahraini R (2015) Genetic trends and parameters of honey production, swarming and defense behavior in Iranian honeybee (*Apis mellifera meda*) Colonies. *J Agr Sci Tech* 17:1735–1742
- Tarpy DR, Nielsen DI (2002) Sampling error, effective paternity, and estimating the genetic structure of honey bee colonies (Hymenoptera: Apidae). *Ann Entomol Soc Am* 95:513–528
- Tarpy DR, Vanengelsdorp D, Pettis JS (2013) Genetic diversity affects colony survivorship in commercial honey bee colonies. *Naturwissenschaften* 100:723–728
- Wragg D, Marti-Marimon M, Basso B, Bidanel JP, Labarthe E, Bouchez O, Le Conte Y, Vignal A (2016) Whole-genome resequencing of honeybee drones to detect genomic selection in a population managed for royal jelly. *Sci Rep* 6:27168. <https://doi.org/10.1038/srep27168>
- Wytrychowski M, Chenavas S, Daniele G, Casabianca H, Batteau M, Guibert S, Brion B (2013) Physicochemical characterisation of French royal jelly: comparison with commercial royal jellies and royal jellies produced through artificial bee-feeding. *J Food Compos Anal* 29(2):126–133
- Zakour MK, Ehrhardt K, Bienefeld K (2012) First estimate of genetic parameters for the Syrian honey bee *Apis mellifera syriaca*. *Apidologie* 43:634–642. <https://doi.org/10.1007/s13592-012-0130-4>

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