the **genetics**society

ARTICLE

Short-term effects of controlled mating and selection on the genetic variance of honeybee populations

Manuel Du [1] · Richard Bernstein[1] · Andreas Hoppe[1] · Kaspar Bienefeld[1]

Abstract

Directional selection in a population yields reduced genetic variance due to the Bulmer effect. While this effect has been thoroughly investigated in mammals, it is poorly studied in social insects with biological peculiarities such as haplo-diploidy or the collective expression of traits. In addition to the natural adaptation to climate change, parasites, and pesticides, honeybees increasingly experience artificial selection pressure through modern breeding programs. Besides selection, many honeybee breeding schemes introduce controlled mating. We investigated which individual effects selection and controlled mating have on genetic variance. We derived formulas to describe short-term changes of genetic variance in honeybee populations and conducted computer simulations to confirm them. Thereby, we found that the changes in genetic variance depend on whether the variance is measured between queens (inheritance criterion), worker groups (selection criterion), or both (performance criterion). All three criteria showed reduced genetic variance under selection. In the selection and performance criteria, our formulas and simulations showed an increased genetic variance through controlled mating. This newly described effect counterbalanced and occasionally outweighed the Bulmer effect. It could not be observed in the inheritance criterion. A good understanding of the different notions of genetic variance in honeybees, therefore, appears crucial to interpreting population parameters correctly.

Introduction

Many populations of social insects are currently exposed to intense levels of natural selection due to climate change, pesticides, habitat fragmentation, or parasites (Chapman and Bourke 2001; Le Conte and Navajas 2008; Mikheyev et al. 2015). These factors greatly influence the genetic structure of these populations. For many populations of the European honeybee (*Apis mellifera*), additional genetic changes are to be expected due to artificial selection, since great efforts have been undertaken to establish modern breeding programs for honeybees in Europe (Büchler et al. 2013; Uzunov et al. 2015). It is known that intensified

selection leads to a rapid reduction of genetic variance due to a disequilibrium effect which was first described by Bulmer (1971). This Bulmer effect appears for both natural (Hawthorne 1998) and artificial selection (Atkins and Thompson 1986). For diploid dioecious species, it can be explained as follows.

Assume a purely additive heritable trait that follows the assumptions of Fisher's infinitesimal model (Fisher 1918). Further, assume a population with discrete generations in which every individual z has a true-breeding value u_z with respect to that trait. Under panmixia and the absence of selection (assuming that effects of inbreeding, mutation, and genetic drift are negligible), the variance of breeding values within one generation g of the population is precisely the additive genetic variance σ_A^2 of the trait:

$$\text{var}(u_z | z \text{ is in generation } g) = \sigma_A^2. \tag{1}$$

The breeding value of an individual z is inherited from its sire $s(z)$ and dam $d(z)$ via the formula

$$u_z = \frac{1}{2} \cdot u_{s(z)} + \frac{1}{2} \cdot u_{d(z)} + \delta_{s(z)} + \delta_{d(z)}, \tag{2}$$

✉ Manuel Du
   manuel.du@hu-berlin.de

[1] Institute for Bee Research Hohen Neuendorf, Hohen Neuendorf, Germany

where the Mendelian sampling terms $\delta_{s(z)}$ and $\delta_{d(z)}$ are independent of $u_{s(z)}$ and $u_{d(z)}$ and one another, and each has variance $\frac{1}{4} \cdot \sigma_A^2$. Passing to variances in Eq. (2) thus yields

$$
\begin{aligned}
&\text{var}(u_z | z \text{ is in generation } g+1) \\
&= \tfrac{1}{4} \cdot \text{var}(u_s | s \text{ is sire in generation } g) \\
&\quad + \tfrac{1}{4} \cdot \text{var}(u_d | d \text{ is dam in generation } g) \\
&\quad + \tfrac{1}{2} \cdot \text{cov}(u_s, u_d | s, d \text{ are parents in generation } g) \\
&\quad + \tfrac{1}{2} \cdot \sigma_A^2.
\end{aligned}
\tag{3}
$$

In the absence of selection or assortative mating, the dams and sires in generation g are independent random samples of the population, whence their variances are again σ_A^2 and their covariance is 0. In this way, Eq. (3) illustrates how in ideal populations the additive genetic variance remains constant over generations. In the presence of selection, however, while the variances of the Mendelian sampling terms $\delta_{s(z)}$ and $\delta_{d(z)}$ remain unchanged, the sires and dams in generation g are no longer a representative sample of the population and their variance will differ. If the genetic variance within the selected sires and dams in generation g are $(1 - \kappa_s) \cdot \sigma_A^2$ and $(1 - \kappa_d) \cdot \sigma_A^2$, respectively, the variance in generation $g + 1$ will be

$$
\text{var}(u_z | z \text{ is in generation } g+1) = \left(1 - \frac{\kappa_s + \kappa_d}{4}\right) \cdot \sigma_A^2.
\tag{4}
$$

As κ_s and κ_d are usually positive, Eq. (4) means a reduction of genetic variance within the population as compared to the case without selection. However, it should be noted that this reduction of variance (unlike reductions caused by drift or inbreeding) is reversible. Once the selection regime ends, the original variance will soon be restored (Bulmer 1971). Similar considerations can be made if a non-zero covariance between parental breeding values is imposed by assortative mating (Tallis and Leppard 1987). If selection is performed as truncation selection with estimated breeding values on a normally distributed trait, κ_j ($= \kappa_s$ or κ_d) can be quantified as

$$
\kappa_j = i_j (i_j - x_j) \cdot \rho_j^2,
\tag{5}
$$

where i_j denotes the sex-specific selection intensity, x_j the standardized truncation point, and ρ_j the correlation between estimated and true breeding values of sires and dams, respectively (Dekkers 1992). For shorter notation, the term $i_j(i_j - x_j)$ is commonly abbreviated by k_j (van der Werf and de Boer 1990).

The biology of social insects, and in particular of honeybees, varies from other species in a number of ways which prevent the theory explained above from carrying over immediately. The most notable differences are the haplo-diploid genetics, the mating behavior, and the expression of selection traits in a colony rather than in individuals. Honeybees separate into three castes: queens, workers, and drones. While the male drones are haploid and develop from unfertilized eggs, queens and workers are female and diploid; however, in general, only drones and queens can reproduce. A colony consists of one queen and up to 40,000 workers, all of which are daughters of the queen. Depending on the season, the queen can further produce several hundreds of drones. All practical work in- and outside the hive is carried out by workers and queen, often in close collaboration, whence most economic traits in honeybees are assumed to have maternal (queen) and direct (worker group) effects (Bienefeld and Pirchner 1991; Brascamp et al. 2016; Chevalet and Cornuet 1982). While queens and drones are generally seen as individuals, workers are usually only regarded as a collective. Shortly after hatching, a newly-born queen undertakes her nuptial flight during which she mates in the air with drones from other colonies. The semen stored during this flight is used later on to fertilize eggs to develop into female offspring. The mating procedure is usually not observable, and artificial selection on the paternal path is therefore not straightforward and often abstained from.

One possibility to enable selection on the paternal path in breeding schemes is isolated mating stations. These consist of secluded geographic areas, like valleys or islands, which are void of honeybee colonies with the exception of a group of colonies that are held there for the purpose of producing drones. Usually, the queens of these drone producing colonies share a common dam, which in turn has been selected for her superior breeding values. The mating stations fulfill two purposes: (a) ensure that only drones with favorable genetic features are involved in the mating process, and (b) provide pedigree information on the paternal side and therefore lead to more accurate breeding value estimations. In honeybee pedigrees, mating stations correspond to sires in pedigrees of other farm animals. Thus, mating stations are also referred to as *pseudo sires* (Bienefeld et al. 2007, 1989; Brascamp and Bijma 2014; Plate et al. 2019a).

Simulation studies have shown that breeding schemes with controlled mating are superior to those without (Plate et al. 2019b). Ideally, the introduction of a new breeding scheme for a honeybee population, therefore, involves both selection and controlled mating for one or several traits and establishing controlled mating (Uzunov et al. 2017). However, there are cases, where selection is performed without controlled mating (Andonov et al. 2019; Pernal et al. 2012; Spivak and Reuter 2001). In the context of conservation programs, situations are also conceivable where mating is controlled but no directional selection is imposed.

Both controlled mating and selection may have short-term implications on the variance structure within the honeybee population which to our knowledge have not been analyzed so far. In this article, we give a theoretical derivation of the short-term effects to be expected upon installation of a new honeybee breeding program. Furthermore, we carried out simulation studies with the program BeeSim (Plate et al. 2019a) to verify and quantify our results.

Theory

When it comes to the inheritance of breeding values in honeybees, three different pathways have to be distinguished, depending on whether the recipient is a queen, a drone, or a worker group (Plate et al. 2019a). Throughout, for a queen q, we denote by $Q(q)$ her dam queen and by $D(q)$ her sire drone. Likewise, let $Q(d)$ be the dam queen of a drone d. Finally, for a worker group w, let $Q(w)$ be the dam queen and let $D_1(w), ..., D_{n_D}(w)$ be the sire drones, i.e., the drones which $Q(w)$ mated with. (See Table 1 for an overview of the used variables.) Most breeding traits in honeybees possess maternal and direct effects, which generally have different additive genetic variances $\sigma_{A,m}^2$ and $\sigma_{A,d}^2$, respectively, as well as a covariance $\sigma_{A,md}$, which

usually takes on negative values (Bienefeld and Pirchner 1991; Brascamp et al. 2016). While direct breeding values are only expressed in worker groups and maternal breeding values are only expressed in queens, all individuals possess true breeding values for both effects. To acknowledge this, we denote breeding values of queens, drones, or worker groups ($\mathbf{u}_q$, $\mathbf{u}_d$, and $\overline{\mathbf{u}}_w$) as vectors $\mathbf{u}_z = \begin{bmatrix} u_z^{mat} \\ u_z^{dir} \end{bmatrix}$. Breeding values of worker groups ($\overline{\mathbf{u}}_w$) are equipped with a bar to signify that they are average values of the individual workers of a colony. The additive genetic variance is denoted as a matrix $\mathbf{V}_A = \begin{bmatrix} \sigma_{A,m}^2 & \sigma_{A,md} \\ \sigma_{A,md} & \sigma_{A,d}^2 \end{bmatrix}$. With this notation, we have for a queen q:

$$\mathbf{u}_q = \frac{1}{2} \cdot \mathbf{u}_{Q(q)} + \mathbf{u}_{D(q)} + \boldsymbol{\delta}_{Q(q)}, \tag{6}$$

where the Mendelian sampling term $\boldsymbol{\delta}_{Q(q)}$ has variance $\frac{1}{4} \cdot \mathbf{V}_A$. Note that since drones are haploid, they pass on all of their genetic information and there is no Mendelian sampling among their gametes. For a drone d, we have:

$$\mathbf{u}_d = \frac{1}{2} \cdot \mathbf{u}_{Q(d)} + \boldsymbol{\delta}_{Q(d)}, \tag{7}$$

Table 1 Notation key.

a_{DPQ}	Average relationship between drone producing queens on a mating station
$C_{p,a_{DPQ}}$	$= \frac{1}{4} \cdot (p + (1-p)a_{DPQ})$; genetic covariance of drones on a mating station
$\mathbf{c}_w$	$= cov(\hat{I}_w, \overline{\mathbf{u}}_w)'$; covariance of selection index and workers' breeding values
$D(q)$	Sire drone of a queen q
$D_1(w), ..., D_{n_D}(w)$	Sire drones of a worker group w
$\boldsymbol{\delta}_Q$	Mendelian sampling in the inheritance from a queen Q
h_m^2, h_d^2	Maternal and direct heritability, resp.
$\hat{I}_w$	$= \hat{\overline{u}}_w^{mat} + \hat{\overline{u}}_w^{dir}$; selection index
$i_{\hat{I}_w}$	Selection intensity of index selection
$k_{\hat{I}_w}$	$= i_{\hat{I}_w}(i_{\hat{I}_w} - x_{\hat{I}_w})$; reduction factor of selection index variance under selection
n_D	Number of drones a queen mates with
n_d, n_s	Number of dams of breeding queens and drone producing queens per year, resp.
p	Probability of two drones on a mating station to share the same dam queen
$Q(q), Q(w), Q(d)$	Dam queen of a queen q, worker group w, or drone d, resp.
r_{md}	Genetic correlation between maternal and direct effects
$\sigma_{A,m}^2, \sigma_{A,d}^2$	Maternal and direct additive genetic variance, resp.
$\sigma_{A,md}$	Additive genetic covariance between maternal and direct effects
σ_E^2	Residual variance
$\sigma_{\hat{I}_w}^2$	Variance of selection index
t_{cont}, t_{sel}	Time of introduction of controlled mating and selection, resp.
$\mathbf{u}_q, \overline{\mathbf{u}}_w, \mathbf{u}_d$	True breeding value of a queen q, worker group w, or drone d, resp.
$\hat{\mathbf{u}}_q, \hat{\overline{\mathbf{u}}}_w$	(BLUP)-estimated breeding value of a queen q, or worker group w, resp.
$\mathbf{V}_A$	Additive genetic variance matrix
$x_{\hat{I}_w}$	Standardized truncation point of index selection

where again the Mendelian sampling term $\boldsymbol{\delta}_{Q(d)}$ has variance $\frac{1}{4} \cdot \mathbf{V}_A$. Finally, for a worker group w:

$$\bar{\mathbf{u}}_w = \frac{1}{2} \cdot \mathbf{u}_{Q(w)} + \frac{1}{n_D} \cdot \sum_{i=1}^{n_D} \mathbf{u}_{D_i(w)}. \tag{8}$$

There is no Mendelian sampling term in the inheritance to worker groups because the breeding values of worker groups are means over a large number of individual workers. Equation (8) makes the simplifying assumption that all drones contribute the same relative number of workers, which is not guaranteed in reality.

We start our investigation of the additive genetic variance within a honeybee population considering the queens of that population. Passing to variances in Eq. (6) yields

$$\begin{aligned}
\operatorname{var}(\mathbf{u}_q | q \text{ is queen}) = {}& \tfrac{1}{4} \cdot \operatorname{var}(\mathbf{u}_{Q(q)} | q \text{ is queen}) \\
& + \operatorname{var}(\mathbf{u}_{D(q)} | q \text{ is queen}) \\
& + \operatorname{cov}(\mathbf{u}_{Q(q)}, \mathbf{u}_{D(q)} | q \text{ is queen}) \\
& + \tfrac{1}{4} \cdot \mathbf{V}_A.
\end{aligned} \tag{9}$$

The variance in breeding values for drones is derived from Eq. (7):

$$\operatorname{var}(\mathbf{u}_d | d \text{ is drone}) = \frac{1}{4} \cdot \operatorname{var}(\mathbf{u}_{Q(d)} | d \text{ is drone}) + \frac{1}{4} \cdot \mathbf{V}_A. \tag{10}$$

Turning to the worker groups, we assume for simplicity that every queen mates with the same number n_D of drones. Consequently, passing to variances in Eq. (8) yields

$$\begin{aligned}
& \operatorname{var}(\bar{\mathbf{u}}_w | w \text{ is worker group}) \\
&= \tfrac{1}{4} \cdot \operatorname{var}(\mathbf{u}_{Q(w)} | w \text{ is worker group}) \\
&\quad + \operatorname{cov}(\mathbf{u}_{Q(w)}, \mathbf{u}_{D_i(w)} | w \text{ is worker group}) \\
&\quad + \tfrac{1}{n_D} \cdot \operatorname{var}(\mathbf{u}_{D_i(w)} | w \text{ is worker group}) \\
&\quad + \tfrac{n_D - 1}{n_D} \cdot \operatorname{cov}(\mathbf{u}_{D_i(w)}, \mathbf{u}_{D_j(w)} | w \text{ is worker group}, \ i \neq j).
\end{aligned} \tag{11}$$

Equations (9)–(11) may be assumed for any honeybee population. To draw further conclusions, it is necessary to make additional assumptions on the population structure. In particular, it plays a role if mating is controlled or uncontrolled and whether or not there is selection.

Uncontrolled mating, no selection

We first consider the case that no selection is employed and mating always occurs uncontrolledly. In this case, dam queens and sire drones of queens are independent random

samples of the entire queen and drone population, respectively, whence in Eq. (9), the terms "$\operatorname{var}(\mathbf{u}_{Q(q)} | q \text{ is queen})$" and "$\operatorname{var}(\mathbf{u}_{D(q)} | q \text{ is queen})$" may be replaced by "$\operatorname{var}(\mathbf{u}_q | q \text{ is queen})$" and "$\operatorname{var}(\mathbf{u}_d | d \text{ is drone})$", respectively, and the covariance term vanishes. Likewise, all queens have the same chance to become the dam of a drone, whence the term "$\operatorname{var}(\mathbf{u}_{Q(d)} | d \text{ is drone})$" in Eq. (10) is identical with "$\operatorname{var}(\mathbf{u}_q | q \text{ is queen})$".

Combining the thus modified Eqs. (9) and (10) yields

$$\begin{aligned}
\operatorname{var}(\mathbf{u}_q | q \text{ is queen}) = {}& \tfrac{1}{4} \cdot \operatorname{var}(\mathbf{u}_q | q \text{ is queen}) \\
& + \tfrac{1}{4} \cdot \operatorname{var}(\mathbf{u}_q | q \text{ is queen}) + \tfrac{1}{4} \cdot \mathbf{V}_A + \tfrac{1}{4} \cdot \mathbf{V}_A \\
= {}& \tfrac{1}{2} \cdot \operatorname{var}(\mathbf{u}_q | q \text{ is queen}) + \tfrac{1}{2} \cdot \mathbf{V}_A.
\end{aligned} \tag{12}$$

This shows in analogy to Eq. (3), how the genetic variance in the queen population remains constantly $\mathbf{V}_A$:

$$\operatorname{var}(\mathbf{u}_q | q \text{ is queen}) = \mathbf{V}_A. \tag{13}$$

As a consequence, we obtain from Eq. (10)

$$\operatorname{var}(\mathbf{u}_d | d \text{ is drone}) = \frac{1}{2} \cdot \mathbf{V}_A. \tag{14}$$

Like in the inheritance to queens, the covariance terms also vanish in the inheritance to worker groups (Eq. (11)). Since the sire drones of worker groups form a random sample of all drones, we obtain from Eqs. (13) and (14):

$$\operatorname{var}(\bar{\mathbf{u}}_w | w \text{ is worker group}) = \left(\frac{1}{4} + \frac{1}{2n_D}\right) \cdot \mathbf{V}_A. \tag{15}$$

Controlled mating without selection

The situation that mating on isolated mating stations is imposed but no directional selection is carried out is very rare in honeybee breeding in reality. Nevertheless, in very small populations, like that of the Sicilian honeybee (*A. m. siciliana*), which are endangered by introgression of other subspecies and where the effective population size is too small to allow for directional selection, it may be a viable conservation strategy (Muñoz et al. 2014).

In this case, Eqs. (13) and (14) remain unchanged. The reason is that, while sire drones and drone producing queens are no longer random samples of the population, they still are unselected descendants of the dams of the drone producing queens which in turn are randomly chosen among the population.

But when it comes to the variance structure of worker groups, differences appear in the term $\operatorname{cov}(\mathbf{u}_{D_i(w)}, \mathbf{u}_{D_j(w)})$, as the drones which are involved in a controlled mating are generally related, whence there are positive covariances between their breeding values. For two drones $D_i(w) \neq D_j(w)$ involved in a controlled mating, we may assume a

probability p that they descend from the same drone producing queen and probability $1 - p$ that their dams are different queens on the same mating station. Usually, this probability p is assumed to depend only on the number of drones producing queens on a mating station and potentially the number n_D of drones that mate with a queen (Brascamp and Bijma 2014). In consequence, we have for $D_i(w) \neq D_j(w)$:

$$
\begin{aligned}
&\text{cov}(\mathbf{u}_{D_i(w)}, \mathbf{u}_{D_j(w)}) \\
&= \text{cov}(\tfrac{1}{2} \cdot \mathbf{u}_{Q(D_i(w))} + \boldsymbol{\delta}_{Q(D_i(w))}, \tfrac{1}{2} \cdot \mathbf{u}_{Q(D_j(w))} + \boldsymbol{\delta}_{Q(D_j(w))}) \\
&= \tfrac{1}{4} \cdot \text{cov}(\mathbf{u}_{Q(D_i(w))}, \mathbf{u}_{Q(D_j(w))}) \\
&= \tfrac{1}{4} \cdot (p + (1-p)a_{\text{DPQ}}) \cdot \mathbf{V}_A,
\end{aligned}
\tag{16}
$$

where a_{DPQ} denotes the average relationship between drone producing queens on a mating station. For shorter notation, we will denote this covariance term by

$$
C_{p,a_{\text{DPQ}}} := \frac{1}{4} \cdot (p + (1-p)a_{\text{DPQ}}).
\tag{17}
$$

From Eq. (16), we conclude in the case of controlled mating:

$$
\text{var}(\overline{\mathbf{u}}_w | w \text{ is worker group}) = \left(\frac{1}{4} + \frac{1}{2n_D} + \frac{n_D - 1}{n_D} \cdot C_{p,a_{\text{DPQ}}} \right) \cdot \mathbf{V}_A.
\tag{18}
$$

The values of p and a_{DPQ} depend on assumptions on the composition of drone producing queens on a mating station, the distribution of offspring to drones and queens, and the history of the breeding program (Brascamp and Bijma 2014, 2019). However, for reasonable assumptions (a sister group of eight drone producing queens on a mating station, each queen mates with 12 drones), one can estimate $p \approx 0.125$ (Bienefeld et al. 2007) and $a_{\text{DPQ}} \approx 0.32$ (Brascamp and Bijma 2014, 2019). Assuming these values, we arrive at the estimate that the introduction of controlled mating without selection can increase the variance of true breeding values of worker groups by around 31.8%.

Influence of selection

As in other species, the introduction of selection has an influence on the genetic variance of the population. Truncation selection in honeybees is generally imposed as follows: by best linear unbiased prediction (BLUP) breeding value estimation (Henderson 1975), estimated breeding values $\hat{\mathbf{u}}_q$ and $\hat{\overline{\mathbf{u}}}_w$ for queens and worker groups are determined and fertilized queens are then selected based on the index $\hat{I}_w = \hat{\overline{u}}_w^{\text{mat}} + \hat{\overline{u}}_w^{\text{dir}}$, formed as the sum of the estimated maternal and direct breeding values of their worker groups.

The changes of the variance structure of the population due to selection can be described in terms of three key variables: the selection intensity $i_{\hat{I}_w}$, with which the selection index is selected, the variance $\sigma_{\hat{I}_w}^2$ of the selection index, and the vector of covariances $\mathbf{c}_w = \begin{bmatrix} \text{cov}(\overline{u}_w^{\text{mat}}, \hat{I}_w) \\ \text{cov}(\overline{u}_w^{\text{dir}}, \hat{I}_w) \end{bmatrix}$ between the true maternal and direct breeding values of the worker groups and the selection index.

To derive these dependencies, we build on a general fact about normally distributed random vectors which was first formulated by Pearson (1903) and reviewed in modern notation by e.g., Gianola et al. (1989): let $\mathbf{y}$ and $\mathbf{w}$ be jointly normally distributed characteristics of a population with variance matrix $\text{var}\left(\begin{bmatrix} \mathbf{y} \\ \mathbf{w} \end{bmatrix} \right) = \begin{bmatrix} \mathbf{V}_{yy} & \mathbf{V}_{yw} \\ \mathbf{V}_{wy} & \mathbf{V}_{ww} \end{bmatrix}$. If then, by the selection of some individuals, the variance structure regarding $\mathbf{y}$ is changed to $\mathbf{V}_{yy}^{\text{sel}}$, the variance structure regarding $\mathbf{w}$ can be expected to change to

$$
\mathbf{V}_{ww}^{\text{sel}} = \mathbf{V}_{ww} - \mathbf{V}_{wy} \mathbf{V}_{yy}^{-1} (\mathbf{V}_{yy} - \mathbf{V}_{yy}^{\text{sel}}) \mathbf{V}_{yy}^{-1} \mathbf{V}_{yw}.
\tag{19}
$$

For a queen Q with worker group w, the BLUP-estimated breeding value $\hat{\overline{\mathbf{u}}}_w$ of the worker group coincides with the estimated breeding value $\hat{\mathbf{u}}_q$ of an unphenotyped daughter queen q of Q (Brascamp and Bijma 2019). Likewise, the vector $\mathbf{c}_w$ coincides with the vector of covariances of the maternal and direct true breeding values of q with the selection index $\hat{I}_w$ (proof given in Appendix A). Therefore, under the absence of selection, we have for a worker group w and queen q with $Q(w) = Q(q)$ the joint variance structure

$$
\text{var}\left(\begin{bmatrix} \hat{I}_w \\ \mathbf{u}_q \end{bmatrix} \right) = \begin{bmatrix} \sigma_{\hat{I}_w}^2 & \mathbf{c}_w' \\ \mathbf{c}_w & \mathbf{V}_A \end{bmatrix}.
\tag{20}
$$

Truncation selection with intensity $i_{\hat{I}_w}$ will reduce $\text{var}(\hat{I}_w)$ to

$$
\sigma_{\hat{I}_w, \text{sel}}^2 = i_{\hat{I}_w}(i_{\hat{I}_w} - x_{\hat{I}_w}) \cdot \sigma_{\hat{I}_w}^2,
\tag{21}
$$

where, as in Eq. (5), $x_{\hat{I}_w}$ denotes the standardized truncation point. Consequently, by Eq. (19) and using the abbreviation $k_{\hat{I}_w} := i_{\hat{I}_w}(i_{\hat{I}_w} - x_{\hat{I}_w})$, the variance of $\mathbf{u}_q$ among the next generation of queens is

$$
\text{var}(\mathbf{u}_q)_{\text{sel}} = \mathbf{V}_A - \frac{k_{\hat{I}_w}}{\sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}_w'.
\tag{22}
$$

Equation (22) corresponds to the reduction of genetic variance under index selection in other livestock species as it is written explicitly by e.g., Bijma and Rutten (2002).

In standard theory on the Bulmer effect, the reduction of genetic variance in the population is inferred from the variance among the selected parents (cf. Eq. (4)). For honeybees, we circumvented this by the use of Pearson's formula (19). It is, however, also possible to calculate the genetic variances and covariances among the selected queens and the drones they mated with. We present these calculations in Appendix B. Of the results derived there, is noteworthy that even in the absence of assortative mating there is a nonzero covariance between the true breeding values of selected queens and the drones they mated with. This covariance is negative definite, reflecting that a mediocre queen will only be selected if it is mated with excellent drone material, whereas a queen of exceptional quality may still be selected if she is mated with drones of lower quality.

The consequences of selection for the genetic variances among drones and worker groups are omitted in this manuscript but can be worked out following similar lines as those specified in Appendix B.

Notions of total genetic variance

In order to assess and judge the status of a honeybee population regarding genetic variance in practice, it is necessary to condense the different variances of direct and maternal effects for the members of the three castes into one notion of total additive genetic variance. In practical applications, focus mostly is not laid on individual queens or worker groups; instead, colonies consisting of a queen and her offspring workers are seen as the relevant entities. For these, different combinations of direct and maternal breeding values are considered to be of interest. The sum of direct and maternal breeding values of a queen, $u_q^{\mathrm{mat}} + u_q^{\mathrm{dir}}$, is the so-called inheritance criterion (IC). Its variance structure shows the possibilities of genetic inheritance to future generations (Brascamp and Bijma 2019). The performance criterion (PC) is the sum of the maternal breeding value of a queen and the direct breeding value of her worker group, $u_{Q(w)}^{\mathrm{mat}} + \overline{u}_w^{\mathrm{dir}}$, and describes the genetic contribution to the phenotype of a colony (Plate et al. 2019a). Finally, the sum of direct and maternal breeding values of a worker group, $\overline{u}_w^{\mathrm{mat}} + \overline{u}_w^{\mathrm{dir}}$, is the so-called selection criterion (SC), because the estimation of this value determines whether a queen will be selected for reproduction (Brascamp and Bijma 2014; Plate et al. 2019a).

The theory developed above implies that the introduction of controlled mating increases the variance in the performance and selection criteria (because these consider worker groups), but leaves the IC unaffected. The introduction of selection decreases the genetic variance of a honeybee population in all three criteria (IC, PC, and SC). In the following, we confirm and quantify these theoretical findings with the help of a computer simulation study.

Methods

We used the program BeeSim (Plate et al. 2019a) to simulate a honeybee population consisting of 500 colonies per year over the course of 20 years. All queens, drones, and worker groups were simulated individually and inherited a trait following the infinitesimal model according to Eqs. (6)–(8). For the first years, no selection was carried out, and queens mated uncontrolledly. Uncontrolled mating was realized as in Plate et al. (2019b); i.e., the drones were provided by randomly chosen queens of ages between 1 and 3 years (see Fig. 1A). Starting from year t_{cont}, mating took place on isolated mating stations. From year t_{sel} on, the reproducing queens were selected after a BLUP breeding value estimation like previously simulated in (Plate et al. 2019a, 2020). We simulated three different situations regarding the introduction of controlled mating and selection (see Fig. 2): (a) first controlled mating, then BLUP selection: ($t_{\mathrm{cont}} = 7$, $t_{\mathrm{sel}} = 14$), (b) first selection, then controlled mating ($t_{\mathrm{cont}} = 14$, $t_{\mathrm{sel}} = 7$), and (c) simultaneous start of controlled mating and BLUP selection ($t_{\mathrm{cont}} = t_{\mathrm{sel}} = 10$).

Throughout the simulations, in each year n_d of the 2-year-old queens of that year were chosen to serve as dams for the next generation, producing $500/n_d$ offspring queens each (see Fig. 1). While that choice was random up to year $t_{\mathrm{sel}} - 1$, from year t_{sel} on, the queens were chosen by truncation selection based on BLUP breeding values which were obtained by BLUPF90 (Misztal et al. 2002) with the use of a honeybee specific relationship matrix (Bernstein et al. 2018; Brascamp and Bijma 2014). We simulated the different values of $n_d = 50$, $n_d = 100$, and $n_d = 250$. In the years with BLUP selection, different values of n_d represent different selection intensities on the maternal path.

In the years with controlled mating (from t_{cont} on), n_s of the 3-year-old queens were chosen to serve as dams of the drone producing queens of one mating station each (see Fig. 1B). In case of no selection, these queens were chosen at random, but from year t_{sel} on, they were chosen based on BLUP truncation selection. We simulated the different values $n_s = 5$, $n_s = 25$, and $n_s = 50$ to represent different selection intensities on the paternal path. Each mating station comprised eight drone producing queens. Regardless if mating was controlled or uncontrolled, all queens mated with $n_D = 12$ drones.

Finally, we made two different choices for the selection trait, which differed mainly in the strength of the negative correlation between direct and maternal effects. For both traits, we chose a maternal additive genetic variance of $\sigma_{A,m}^2 = 1$, a direct additive genetic variance of $\sigma_{A,d}^2 = 2$, and a residual variance of $\sigma_E^2 = 1$. Then in one case we chose a moderate negative covariance of $\sigma_{A,md} = -0.5$, and in the other case a stronger covariance of $\sigma_{A,md} = -1$. The former

Fig. 1 Scheme of reproduction. A Scheme of reproduction under uncontrolled mating. Dams of queens are 2 years old and selected either randomly or by BLUP breeding value estimation. Dams of drones are 1–3 years old and are randomly selected. **B** Scheme of reproduction under controlled mating. Dams of queens are 2 years old and grand dams of drones are 3 years old. Both are selected either randomly or by BLUP breeding value estimation. This figure is inspired by Figure 1 in Plate et al. (2019a).
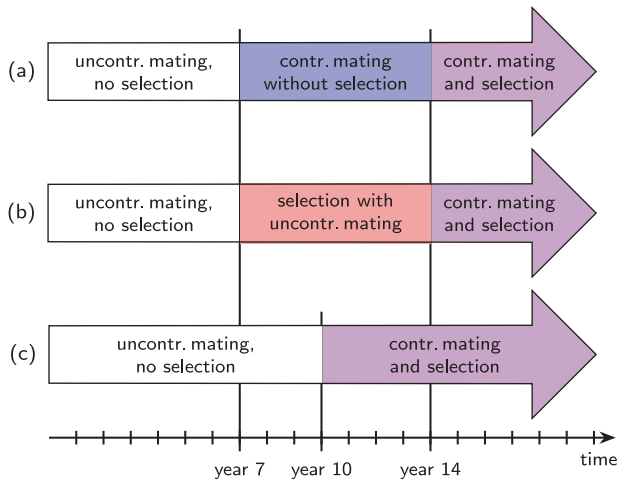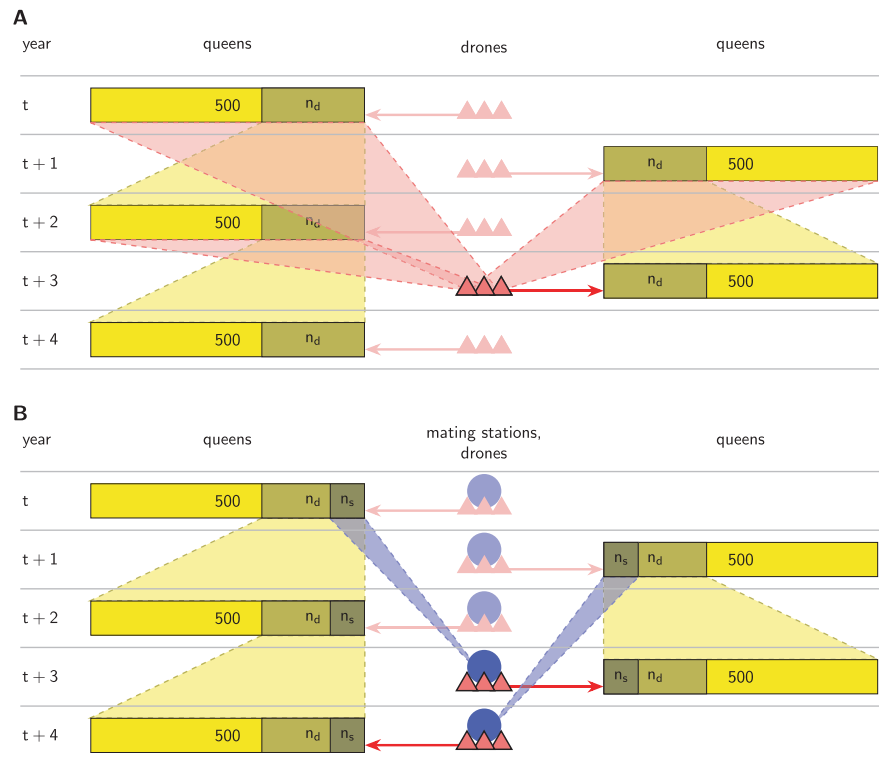

Fig. 2 Breeding schemes. The three different simulated set-ups regarding the order of introduction of controlled mating and selection: (**a**) controlled mating introduced before selection, (**b**) selection introduced before controlled mating, and (**c**) simultaneous introduction of controlled mating and selection.

trait yields a maternal heritability of $h_m^2 = 0.46$, a direct heritability of $h_d^2 = 0.31$, and a correlation between effects of $r_{md} = -0.35$. The corresponding values for the latter trait are $h_m^2 = 0.60$, $h_d^2 = 0.40$, and $r_{md} = -0.71$ (see Brascamp and Bijma (2019) for a detailed description of the calculation of heritabilities for honeybees).

The different combinations of choices for t_{cont}, t_{sel}, n_d, n_s, and $\sigma_{A,md}$ made up a total of 54 different simulation set-ups.

The simulations for each set-up were repeated 200 times for stable results. To assess the changes in genetic variance, we compared the values attained in the year before the introduction of a new breeding strategy with the results from the year of introduction; i.e., genetic variance levels of years $t_{cont} - 1$ and $t_{sel} - 1$ were compared with those of years t_{cont} and t_{sel}, respectively.

Results

The simulations corroborated the theoretical findings. The introduction of controlled mating increased the population variance in the PC and the SC which are (partly) determined by the worker groups but had little effect on the IC which is measured only in the queens. Meanwhile, the introduction of selection caused a reduction of variance in all criteria (see Fig. 3). A similar, though less pronounced, the pattern could be observed for the direct and maternal effects individually (see Supplementary Fig. S1). In the following, we present our findings in greater detail.

Introduction of controlled mating

Independent of the order of introduction of controlled mating and selection, controlled mating did practically not affect the genetic variance in the IC. Upon introduction of controlled mating, reductions or increases in this criterion
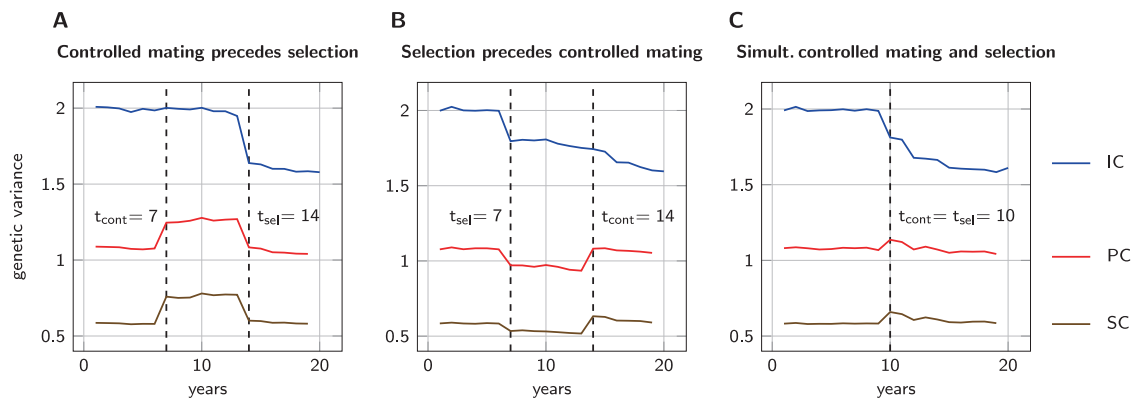
Fig. 3 Genetic variance. Changes of different notions of genetic variance in a honeybee population following the introduction of controlled mating and/or BLUP selection. Results are shown for the parameters $n_d = 100$, $n_s = 25$, and $r_{md} = -0.35$. **A** Controlled mating introduced before selection. **B** Selection introduced before controlled mating. **C** Simultaneous introduction of controlled mating and selection.

Table 2 Change (in %) of population variance after the introduction of controlled mating.

		Controlled mating before selection			Selection before controlled mating		
		50 dams	100 dams	250 dams	50 dams	100 dams	250 dams
$r_{md} = -0.35$	IC	−1.0	+0.1	+0.5	−0.6	−0.6	−0.2
	PC	+13.4	+14.1	+15.0	+14.9	+15.2	+14.1
	SC	+24.7	+26.2	+27.0	+20.5	+20.9	+20.3
$r_{md} = -0.71$	IC	+0.5	−0.4	+0.1	+0.4	−0.7	−0.1
	PC	+27.0	+27.2	+27.0	+26.4	+30.2	+28.6
	SC	+27.4	+26.4	+27.1	+21.8	+24.4	+23.8

Relative change (in %) of genetic variance in the three different criteria IC, PC, and SC, after the introduction of controlled mating. Results are shown for different correlations r_{md} between direct and maternal effects and different numbers of selected dams per year.

were mostly below 1% and did not show any clear direction (see Table 2 and Fig. 3A, B). The increase of genetic variance in the PC due to the introduction of controlled mating depended heavily on the selection trait. For the trait with moderate correlation between effects, $r_{md} = -0.35$, there was an increase between 13.4 and 15.2%; for the trait with strong correlation, $r_{md} = -0.71$, the effect was significantly stronger: between 26.4 and 31.2%. Finally, the SC showed similar behavior in both traits. If controlled mating was introduced before selection, it caused an increase of genetic variance of between 24.7 and 27.4%; if it was introduced after selection the increase of variance was only between 20.3 and 24.4%. In the breeding schemes which introduced controlled mating after selection, the increase of genetic variance in the SC was slightly higher, when the negative correlation between maternal and direct effects was strong, $r_{md} = -0.71$. There was no clear indication of an effect of the number of dam queens on the intensity of the variance increase (see Table 2).

The introduction of controlled mating yielded significantly different results for breeding schemes with only five pseudo sires from those with 25 or 50 pseudo sires.

With only five pseudo sires the increase of genetic variance in PC and SC was significantly lower than in breeding schemes with more pseudo sires. The increase of variance in both criteria was reduced by roughly a fifth with slight variations between the different traits and selection schemes (see Fig. 4).

Introduction of selection

When the selection was introduced after imposing controlled mating ($t_{cont} = 7$, $t_{sel} = 14$), it caused a decrease of variance in the population in all variance criteria (IC, PC, and SC) of between 8.8 and 21.2% (see Table 3 and Fig. 3A). When the introduction of selection preceded the introduction of controlled mating ($t_{cont} = 14$, $t_{sel} = 7$), the reduction of variance appeared less severe, between 3.2 and 10.4% (see also Fig. 3B). The reduction was more pronounced for the trait with a moderate negative correlation between effects, $r_{md} = -0.35$ than for the trait with the stronger negative correlation, $r_{md} = -0.71$. Furthermore, a higher selection intensity (fewer dams selected) resulted in a stronger decrease of variance within the population. When

the selection was introduced after controlled mating, the Bulmer effect showed most strongly in the SC, followed by IC and PC. In the case of $t_{cont} > t_{sel}$, the opposite order could be observed: the Bulmer effect affected the PC strongest and the SC weakest. The number of mating stations, i.e., pseudo sires, n_s, had only a minor effect on the results and there was no clear tendency indicating whether a higher number of pseudo sires caused a stronger or weaker Bulmer effect (data not shown). Furthermore, in the case of selection preceding controlled mating, we observed a significant reduction of genetic variance in the IC from year 15 to year 16 by between 3.2 and 7.6% which was mainly caused by a reduced variance of direct breeding values of queens (see Fig. 3B and Supplementary Fig. S1).

Combined effects

The simultaneous introduction of controlled mating and selection ($t_{sel} = t_{cont} = 10$) resulted in an immediate reduction of genetic variance in the IC of 6.9–11.3% ($r_{md} = -0.35$), and of 3.1–6.2% ($r_{md} = -0.71$), respectively. Two years after the introduction of the breeding scheme, a second drop in genetic variance occurred with a reduction
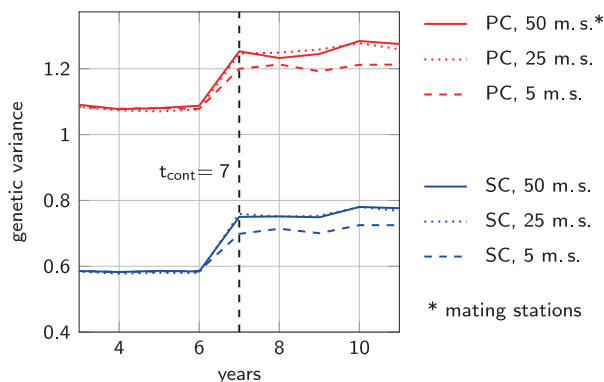

Fig. 4 Influence of the number of pseudo sires on the increase of genetic variance in the PC and SC. Development of genetic variance with the introduction of controlled mating without selection in year $t_{sel} = 7$. Results are shown for $n_d = 100$ and $r_{md} = -0.35$.

of 6.6–8.6% ($r_{md} = -0.35$), and of 3.9–7.2% ($r_{md} = -0.71$), respectively (see Fig. 3C).

In the performance and selection criteria, we observed an initial increase of variance by up to 14.1% ($r_{md} = -0.35$) and 26.2% ($r_{md} = -0.71$), respectively. Subsequently, however, the level of genetic variance in these criteria decreased and thus undermined the initial gain (see Fig. 3C). Comparing the genetic variance of year 9 (directly before the introduction of the breeding scheme) with those of year 13 we could still see an increase of genetic variance for $r_{md} = -0.71$, while for $r_{md} = -0.35$ there were some selection schemes with a positive effect and some with a negative effect on the genetic variance in the PC and SC (see Table 4).

Discussion

Significance of the notions of genetic variance

The different criteria for assessing the genetic variance within a honeybee population (IC, PC, and SC) showed different behavior upon the introduction of breeding schemes, which entails the question of which criterion is the most relevant for honeybee breeding.

Traditionally, the Bulmer effect has mainly been seen as a limitation to the genetic progress a population can reach in total (Dempfle 1974). In this regard, the IC is the most relevant, because only queens can reproduce and thus contribute to the genetic progress of future generations. The IC corresponds to the notion of *heritable variance* in the presence of indirect genetic effects in other agricultural species (Bijma 2011; Brascamp and Bijma 2019). As we have shown, the population variance in this criterion was reduced by the introduction of selection and largely unaffected by the introduction of controlled mating. Therefore, we assume that when new breeding schemes are introduced in honeybees, the Bulmer effect has mainly the same consequences for the breeding success as in other agricultural species.

The other two notions of genetic variance (PC and SC) have more indirect effects on breeding success. The PC

Table 3 Change (in %) of population variance through the introduction of selection.

		Controlled mating before selection			Selection before controlled mating		
		50 dams	100 dams	250 dams	50 dams	100 dams	250 dams
$r_{md} = -0.35$	IC	−18.4	−16.4	−14.3	−9.9	−9.5	−7.6
	PC	−16.6	−14.6	−13.6	−10.4	−10.2	−8.1
	SC	−22.6	−21.2	−20.7	−8.3	−8.0	−6.6
$r_{md} = -0.71$	IC	−13.9	−11.7	−9.0	−5.2	−4.7	−3.7
	PC	−13.6	−10.2	−8.8	−6.8	−6.0	−5.2
	SC	−18.7	−16.3	−14.4	−4.3	−4.0	−3.2

Reduction (in %) of genetic variance in the three different criteria IC, PC, and SC, due to the Bulmer effect upon introduction of BLUP selection. Results are shown for different correlations r_{md} between direct and maternal effects and different numbers of selected dams per year.

Table 4 Change (in %) of population variance by the simultaneous introduction of controlled mating and selection.

		$r_{md} = -0.35$			$r_{md} = -0.71$		
		IC	PC	SC	IC	PC	SC
5 pseudo sires	50 dams	−17.5	−6.1	−2.0	−12.8	+11.1	+7.4
	100 dams	−16.3	−2.1	+2.2	−12.0	+12.4	+6.4
	250 dams	−14.7	−1.8	+1.2	−10.1	+12.2	+8.3
25 pseudo sires	50 dams	−18.0	−2.6	+3.9	−11.8	+16.6	+12.5
	100 dams	−15.9	+2.1	+7.0	−10.4	+22.4	+15.2
	250 dams	−14.1	+1.5	+7.3	−8.0	+21.4	+17.5
50 pseudo sires	50 dams	−18.1	−1.7	+4.2	−11.8	+19.8	+14.9
	100 dams	−16.4	+1.1	+6.6	−10.0	+23.3	+16.5
	250 dams	−13.4	+3.3	+9.5	−7.9	+26.9	+19.3

Relative change (in %) of genetic variance from year 9 to year 13, when both controlled mating and selection were introduced in year 10.

influences the phenotypic variance of a population. An increase of genetic variance in the PC without an increase of variance in the IC does not increase the evolutionary potential of a population. However, in a study on microbia, van Boxtel et al. (2017) showed that non-inheritable phenotypic variance can be powerful insurance against sudden extinction of (sub-) populations. Breeding decisions in honeybees are based on the SC. Increased variance in this criterion, like it is caused by the introduction of controlled mating, allows BLUP to yield estimated breeding values with higher accuracy. Consequently, the higher genetic variance in the SC is a minor contributor to the superiority of breeding schemes with controlled mating over those without (Plate et al. 2019b).

Changes of genetic variance

Overall, the simulation results confirmed the theoretical predictions, i.e., that the introduction of selection decreases the genetic variance within the population in all criteria (IC, PC, and SC), while the introduction of controlled mating results in an increase of genetic variance in those criteria which account for worker groups (PC and SC) but does not influence the variance structure of the queens in the population (IC). In the context of honeybees, possible consequences of the Bulmer effect have been mentioned by Moritz (1986), but only under the a priori assumption that it behaves like in other agricultural species. To our knowledge, this is the first thorough investigation of the Bulmer effect for honeybees. In addition, the positive effect of controlled mating on the variance in the PC and SC is described for the first time.

Introduction of controlled mating

Both theory and simulation studies agreed that the introduction of controlled mating increases the genetic variance in the PC and SC but not in the IC. In the situation where controlled mating is imposed on unselected populations, theory and simulations for the SC also confirm each other quantitatively, with the realized increase of genetic variance being only slightly lower than the 31.8% we derived in the theory section.

In selected populations, the increase of genetic variance in the PC and SC in the course of the introduction of controlled mating was lower than in unselected populations (see Table 2 and Fig. 3A, B), because the variance among the dams of the drone producing queens was reduced. Plate et al. (2019b) showed that controlled mating improves the accuracy of breeding value estimation for honeybees. The enhanced accuracy of breeding values after the introduction of controlled mating leads to higher covariances c_w between the selection index and the true breeding values and thus to a further reduction of genetic variance among the dams of future generations according to Eq. (22). This explains the continued reduction of genetic variance over several years after the introduction of controlled mating which was especially pronounced when selection and controlled mating were introduced simultaneously (see Fig. 3C).

The effect of increased variance between worker groups due to controlled mating is to some extent comparable to the effect of assortative mating in other agricultural species as described by Tallis and Leppard (1987, 1988). If parents are mated assortatively, the positive correlation between their breeding values results in an increased genetic variance. In the case of controlled mating of honeybees, the increased genetic variance between worker groups is likewise caused by a positive correlation of the parental breeding values (albeit not between the breeding values of sire and dam but between the multiple sire drones of a worker group).

Another possible interpretation of the described effect is that of a shift of genetic variance. While the genetic variance between worker groups is increased, controlled mating decreases the intracolonial genetic variance (Oldroyd et al. 1992). In fact, much as it is the case for queens, the genetic variance between all individual worker bees in the population remains unaffected by controlled mating.

Introduction of selection

The reduction of genetic variance in the population after the start of selection was significantly stronger when mating was controlled (see Table 3 and Fig. 3A, B). This has two reasons: Controlled mating allows (a) for selection on the paternal path which causes a reduction of variance among the sire drones, and (b) for a more accurate breeding value estimation and thus a further reduced variance among dam queens (Plate et al. 2019b). The different rates of variance reduction for the two traits can be explained by the different accuracies of BLUP selection. Stronger negative covariances between effects have been shown to lead to lower accuracy in the breeding value estimation of honeybees and other agricultural species (Plate et al. 2019a, b; Roehe and Kennedy 1993). In consequence, there is more genetic variance among the selected dams (and in the case of controlled mating also sires) and thus a less pronounced Bulmer effect. Similarly, truncation selection schemes with fewer selected dams show a greater similarity among the selected individuals and thus a stronger Bulmer effect. The number of pseudo sires has a twofold effect. On the one hand, a small number of pseudo sires means a high intensity of selection (large $k_{\hat{I}_w}$), but on the other hand, Plate et al. (2019b) have shown that small numbers of mating stations also reduce the accuracy of the BLUP breeding value estimation (small values in $\mathbf{c}_w$). Equation (22) indicates that these two aspects have opposing effects on the genetic variance. The effect of the number of pseudo sires on the Bulmer effect due to the introduction of selection is therefore involved and depends on cofactors such as genetic parameters and the number of selected dams.

Limitations of theory

Inhomogeneous populations

In our simulations, we assumed a very homogeneous population in which the population size and the number of selected dams remained constant each year, and each dam mated with the same number n_D of drones. In particular, the last assumption is unrealistic unless all queens are artificially inseminated. This is slightly problematic since the putatively constant number n_D is explicitly used in the derivation of the formulas for the variance structure of worker groups (Eq. (11) and derived equations). A natural approach for populations in which queens mate with different numbers of drones would be to use the average number of drones mated to a queen, $\overline{n}_D$, for which estimates have been published (Tarpy and Nielsen 2002). While this is probably acceptable in most situations, it should be mentioned that the replacement of n_D by $\overline{n}_D$ in formulas derived from Eq. (11) will systematically underestimate the genetic variance among worker groups in heterogeneous populations. This is a consequence of Jensen's inequality (Jensen 1906) since the inversion of n_D in Eq. (11) is a convex function.

Small numbers of dams and sires

We assumed that the genetic variance within an unselected sample of the population equals the population variance. In fact, the expected sample variance in a sample of size n is lower than the population variance by a factor $\frac{n-1}{n}$ (Kenney and Keeping 1957). The derived formulas should therefore be used with care if there are only a few pseudo sires or if very few dam queens are selected. In our simulations, an effect of reduced sample variance due to the small sample size could be observed in simulations with only five pseudo sires (Fig. 4). The lower increase of genetic variance in the PC and SC in these cases can be explained by the reduced variance within the sample of the five dams of the drone producing queens on mating stations.

Middle to long term effects

Besides the Bulmer effect, changes in genetic variance can also be caused by genetic drift, inbreeding, and mutation. The effects of these factors on honeybees have been investigated in various studies (Beye et al. 2006; Plate et al. 2019a, 2020; Zayed and Packer 2005). However, they will influence the genetic variance only over an extended period of time and not in the short time frame considered in the present study. The initial gain of genetic variance in the performance and SC after the introduction of controlled mating could therefore become reduced over time as the selection of pseudo sires increases the risks of accumulated inbreeding and genetic drift (Plate et al. 2019a, b, 2020). Furthermore, we mainly quantified the effects of controlled mating and selection to the immediate next generation. In other livestock species, it has been shown that the Bulmer effect can further decrease the genetic variance over several generations until an equilibrium is reached (Van Grevenhof et al. 2012). But following a geometric series, convergence to equilibrium is generally fast and Fig. 3A, B suggests that no large effect occurs in later generations since the curves turn flat immediately after the disruptions in years t_{cont} and t_{sel}.

Genetic model

Like the original work of Bulmer (1971), this study assumes throughout that selection traits follow the infinitesimal model, which is unlikely in reality. Plate et al. (2019a) showed that the choice of genetic model is of little relevance for the short-term genetic gain in simulated honeybee breeding systems. However, this may be different in the context of genetic variance. Turelli and Barton (1990)

derived that short-term changes in genetic variance, such as the Bulmer effect, may behave differently if the selection trait deviates from normality or in the presence of linkage disequilibrium in the unselected population. The high recombination rate for the honeybee genome (Beye et al. 2006) may alleviate such effects.

Other selection schemes

Within-family selection

In our simulations, we used a truncation selection scheme based on BLUP breeding values. For other agricultural species, it has been shown that the reduction of genetic variance due to the Bulmer effect is greatly reduced or entirely absent if within-family selection is applied (Dempfle 1974; Wei et al. 1996). These results can be expected to hold true for honeybees as far as the reduction of genetic variance due to selection is concerned.

In reality, honeybee breeding programs usually realize neither pure truncation selection nor pure within-family selection but a complicated blend of selection systems shaped by heterogeneous decisions of breeders. Furthermore, some breeding systems are open for the introduction of queens with unregistered dams which also affects the variance structure (Brascamp et al. 2016).

Artificial insemination

Artificial insemination of queens is a popular alternative to isolated mating stations (Cobey et al. 2013). In many cases, artificial insemination mimics the situation on a mating station, i.e., drones for the insemination process are selected among colonies whose queens share a common dam. However, differing procedures are possible and sometimes applied in reality. Instead of using drones from several colonies, insemination of a queen can be performed with drones from a single colony (Gerula et al. 2014) or even with a single drone (Harbo 1999). If each queen is inseminated with drones from a single colony, the covariance terms in Eq. (11) will be increased and the increase of genetic variance in the PC and SC due to the controlled mating will be stronger. It will further intensify if the number of drones, n_D, is reduced because this variable is divided by Eq. (11). In a hypothetical population in which all queens are inseminated by randomly selected single drones, Eq. (11) would simplify to

$$\mathrm{var}(\bar{\mathbf{u}}_w | w \text{ is worker group}) = \frac{3}{4} \cdot \mathbf{V}_A. \qquad (23)$$

As another extreme, queens can also be inseminated with semen from very large pools of unrelated drones (Pieplow et al. 2017). If no selection is imposed on the drone sires, Eq. (15) can be used to describe the variance structure among the colonies. However, since the number of drones is very high, n_D can be assumed to tend to infinity, reducing Eq. (15) to

$$\mathrm{var}(\bar{\mathbf{u}}_w | w \text{ is worker group}) = \frac{1}{4} \cdot \mathbf{V}_A. \qquad (24)$$

Equations (23) and (24) describe the maximum and minimum genetic variance among worker groups which can be attained through different insemination schemes (not taking genetic drift, inbreeding, or mutations into account).

Data availability

The dataset used and analyzed in this article is available at https://doi.org/10.5061/dryad.pzgmsbck4. The source code of the simulation program BeeSim is available at https://doi.org/10.5061/dryad.1nh544n.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

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

Appendix A

Covariance of selection index and breeding values of replacement queens

Let w be a worker group and q be an unphenotyped queen with $Q(w) = Q(q)$. Let $\bar{\mathbf{u}}_w$, $\mathbf{u}_q$ and $\hat{\bar{\mathbf{u}}}_w$, $\hat{\mathbf{u}}_q$ be their respective true and BLUP-estimated breeding values. In this appendix,

we will show that

$$\mathrm{cov}(\overline{\mathbf{u}}_w, \hat{\overline{u}}_w^{\mathrm{mat}} + \hat{\overline{u}}_w^{\mathrm{dir}}) = \mathrm{cov}(\mathbf{u}_q, \hat{\overline{u}}_w^{\mathrm{mat}} + \hat{\overline{u}}_w^{\mathrm{dir}}). \tag{25}$$

Proof It is a well-established fact that the estimated breeding values for w and q coincide (Brascamp and Bijma 2019, Equation (8)):

$$\hat{\overline{\mathbf{u}}}_w = \hat{\mathbf{u}}_q. \tag{26}$$

Furthermore, it is a property of every BLUP-estimate $\hat{\mathbf{u}}$ of a true-breeding value $\mathbf{u}$ that (Henderson 1975)

$$\mathrm{cov}(\mathbf{u}, \hat{\mathbf{u}}) = \mathrm{var}(\hat{\mathbf{u}}). \tag{27}$$

$$\mathrm{var}(\mathbf{u}) = \begin{bmatrix} \mathrm{var}(\overline{\mathbf{u}}_w) & \mathrm{cov}(\overline{\mathbf{u}}_w, \mathbf{u}_{Q(w)}) & \mathrm{cov}(\overline{\mathbf{u}}_w, \mathbf{u}_{D_1(w)}) & \cdots & \mathrm{cov}(\overline{\mathbf{u}}_w, \mathbf{u}_{D_{n_D}(w)}) \\ \mathrm{cov}(\overline{\mathbf{u}}_w, \mathbf{u}_{Q(w)}) & \mathrm{var}(\mathbf{u}_{Q(w)}) & \mathrm{cov}(\mathbf{u}_{Q(w)}, \mathbf{u}_{D_1(w)}) & \cdots & \mathrm{cov}(\mathbf{u}_{Q(w)}, \mathbf{u}_{D_{n_D}(w)}) \\ \mathrm{cov}(\overline{\mathbf{u}}_w, \mathbf{u}_{D_1(w)}) & \mathrm{cov}(\mathbf{u}_{Q(w)}, \mathbf{u}_{D_1(w)}) & \mathrm{var}(\mathbf{u}_{D_1(w)}) & \cdots & \mathrm{cov}(\mathbf{u}_{D_1(w)}, \mathbf{u}_{D_{n_D}(w)}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(\overline{\mathbf{u}}_w, \mathbf{u}_{D_{n_D}(w)}) & \mathrm{cov}(\mathbf{u}_{Q(w)}, \mathbf{u}_{D_{n_D}(w)}) & \mathrm{cov}(\mathbf{u}_{D_1(w)}, \mathbf{u}_{D_{n_D}(w)}) & \cdots & \mathrm{var}(\mathbf{u}_{D_{n_D}(w)}) \end{bmatrix}. \tag{30}$$

Therefore:

$$\mathrm{cov}(\overline{\mathbf{u}}_w, \hat{\overline{u}}_w^{\mathrm{mat}} + \hat{\overline{u}}_w^{\mathrm{dir}}) = \mathrm{cov}(\overline{\mathbf{u}}_w, \hat{\overline{\mathbf{u}}}_w) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \overset{(27)}{=} \mathrm{var}(\hat{\overline{\mathbf{u}}}_w) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \overset{(26)}{=} \mathrm{var}(\hat{\mathbf{u}}_q) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\overset{(27)}{=} \mathrm{cov}(\mathbf{u}_q, \hat{\mathbf{u}}_q) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \overset{(26)}{=} \mathrm{cov}(\mathbf{u}_q, \hat{\overline{\mathbf{u}}}_w) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathrm{cov}(\mathbf{u}_q, \hat{\overline{u}}_w^{\mathrm{mat}} + \hat{\overline{u}}_w^{\mathrm{dir}}).$$

q.e.d.

Appendix B

Genetic (co-)variances among selected queens and drones

In this appendix, we expound how BLUP-selection in honeybees affects the genetic (co-)variances among the selected queens and the drones they mated with. The notation is the same as in the main article.

In the absence of selection, the selection index $\hat{I}_w$ and the true breeding values of worker groups $\overline{\mathbf{u}}_w$ have the joint variance

$$\mathrm{var}\left(\begin{bmatrix} \hat{I}_w \\ \overline{\mathbf{u}}_w \end{bmatrix}\right) = \begin{bmatrix} \sigma_{\hat{I}_w}^2 & \mathbf{c}_w' \\ \mathbf{c}_w & W \cdot \mathbf{V}_A \end{bmatrix}, \tag{28}$$

where $W = W_u := \frac{1}{4} + \frac{1}{2n_D}$ if mating is uncontrolled (Eq. (15)) and $W = W_c := \frac{1}{4} + \frac{1}{2n_D} + \frac{n_D - 1}{n_D} \cdot C_{p,a_{\mathrm{DPQ}}}$ if mating is controlled (Eq. (18)). Truncation selection will reduce $\mathrm{var}(\hat{I}_w)$ to $k_{\hat{I}_w} \cdot \sigma_{\hat{I}_w}^2$, whence by Pearson's Eq. (19), the genetic variance among the worker groups of selected queens is

$$\mathrm{var}(\overline{\mathbf{u}}_w)_{\mathrm{sel}} = W \cdot \mathbf{V}_A - \frac{k_{\hat{I}_w}}{\sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}_w'. \tag{29}$$

To conclude the changes of (co-)variances among the true breeding values of selected queens and the drones they mated with, we apply Eq. (19) to the joint variance matrix of the vector $\mathbf{u} := \begin{bmatrix} \overline{\mathbf{u}}_w' & \mathbf{u}_{Q(w)}' & \mathbf{u}_{D_1(w)}' & \cdots & \mathbf{u}_{D_{n_D}(w)}' \end{bmatrix}'$:

Under uncontrolled mating conditions, we have in the absence of selection

$$\begin{aligned} \mathrm{cov}(\overline{\mathbf{u}}_w, \mathbf{u}_{Q(w)}) &= \tfrac{1}{2} \cdot \mathbf{V}_A, \\ \mathrm{cov}(\overline{\mathbf{u}}_w, \mathbf{u}_{D_i(w)}) &= \tfrac{1}{2n_D} \cdot \mathbf{V}_A, \\ \mathrm{var}(\mathbf{u}_{Q(w)}) &= \mathbf{V}_A, \\ \mathrm{cov}(\mathbf{u}_{Q(w)}, \mathbf{u}_{D_i(w)}) &= \mathbf{0}, \\ \mathrm{var}(\mathbf{u}_{D_i(w)}) &= \tfrac{1}{2} \cdot \mathbf{V}_A, \\ \mathrm{cov}(\mathbf{u}_{D_i(w)}, \mathbf{u}_{D_j(w)}) &= \mathbf{0}. \end{aligned} \tag{31}$$

Inserting these values into Eq. (30) and applying the fact that the variance of true breeding values among the worker groups of selected queens is $\mathrm{var}(\overline{\mathbf{u}}_w)_{\mathrm{sel}} = W_u \cdot \mathbf{V}_A - \frac{k_{\hat{I}_w}}{\sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}_w'$ with $W_u = \frac{1}{4} + \frac{1}{2n_D}$ (Eq. (29)), we deduce by using Pearson's Eq. (19):

$$\begin{aligned} \mathrm{var}(\mathbf{u}_{Q(w)})_{\mathrm{sel}} &= \mathbf{V}_A - \frac{k_{\hat{I}_w}}{4W_u^2 \cdot \sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}_w', \\ \mathrm{cov}(\mathbf{u}_{Q(w)}, \mathbf{u}_{D_i(w)})_{\mathrm{sel}} &= -\frac{k_{\hat{I}_w}}{4W_u^2 \cdot n_D \cdot \sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}_w', \\ \mathrm{var}(\mathbf{u}_{D_i(w)})_{\mathrm{sel}} &= \tfrac{1}{2} \cdot \mathbf{V}_A - \frac{k_{\hat{I}_w}}{4W_u^2 \cdot n_D^2 \cdot \sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}_w', \\ \mathrm{cov}(\mathbf{u}_{D_i(w)}, \mathbf{u}_{D_j(w)})_{\mathrm{sel}} &= -\frac{k_{\hat{I}_w}}{4W_u^2 \cdot n_D^2 \cdot \sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}_w'. \end{aligned} \tag{32}$$

Under controlled mating conditions, we have $W_c = \frac{1}{4} + \frac{1}{2n_D} + \frac{n_D-1}{n_D} \cdot C_{p,a_{DPQ}}$, and the analog to the set of Eq. (31) is

$$
\begin{aligned}
\text{cov}(\overline{\mathbf{u}}_w, \mathbf{u}_{Q(w)}) &= \tfrac{1}{2} \cdot \mathbf{V}_A, \\
\text{cov}(\overline{\mathbf{u}}_w, \mathbf{u}_{D_i(w)}) &= \left(W_c - \tfrac{1}{4}\right) \cdot \mathbf{V}_A, \\
\text{var}(\mathbf{u}_{Q(w)}) &= \mathbf{V}_A, \\
\text{cov}(\mathbf{u}_{Q(w)}, \mathbf{u}_{D_i(w)}) &= \mathbf{0}, \\
\text{var}(\mathbf{u}_{D_i(w)}) &= \tfrac{1}{2} \cdot \mathbf{V}_A, \\
\text{cov}(\mathbf{u}_{D_i(w)}, \mathbf{u}_{D_j(w)}) &= C_{p,a_{DPQ}} \cdot \mathbf{V}_A,
\end{aligned}
\tag{33}
$$

and with $\text{var}(\overline{\mathbf{u}}_w)_{\text{sel}} = W_c \cdot \mathbf{V}_A - \frac{k_{\hat{I}_w}}{\sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}'_w$, this yields

$$
\begin{aligned}
\text{var}(\mathbf{u}_{Q(w)})_{\text{sel}} &= \mathbf{V}_A - \frac{k_{\hat{I}_w}}{4W_c^2 \cdot \sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}'_w, \\
\text{cov}(\mathbf{u}_{Q(w)}, \mathbf{u}_{D_i(w)})_{\text{sel}} &= -\frac{(4W_c-1) \cdot k_{\hat{I}_w}}{8W_c^2 \cdot \sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}'_w, \\
\text{var}(\mathbf{u}_{D_i(w)})_{\text{sel}} &= \tfrac{1}{2} \cdot \mathbf{V}_A - \frac{(4W_c-1)^2 \cdot k_{\hat{I}_w}}{16W_c^2 \cdot \sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}'_w, \\
\text{cov}(\mathbf{u}_{D_i(w)}, \mathbf{u}_{D_j(w)})_{\text{sel}} &= C_{p,a_{DPQ}} \cdot \mathbf{V}_A - \frac{(4W_c-1)^2 \cdot k_{\hat{I}_w}}{16W_c^2 \cdot \sigma_{\hat{I}_w}^2} \cdot \mathbf{c}_w \mathbf{c}'_w.
\end{aligned}
\tag{34}
$$

As an alternative to the derivation in the main article, the genetic variance among the next generation of queens (Eq. (22)) can also be obtained by inserting Eqs. (32) or (34) into Eq. (9).

References

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 20:78

Atkins KD, Thompson R (1986) Predicted and realized responses to selection for an index of bone length and body weight in Scottish Blackface sheep. 1. Responses in the index and component traits. Anim Prod 44:421–435

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

Beye M, Gattermeier I, Hasselmann M, Gempe T, Schioett M, Baines JF et al. (2006) Exceptionally high levels of recombination across the honey bee genome. Genome Res 16:1139–1144

Bienefeld K, Ehrhardt K, Reinhardt F (2007) Genetic evaluation in the honey bee considering queen and worker effects—a BLUP—animal model approach. Apidologie 38:77–85

Bienefeld K, Pirchner F (1991) Genetic correlations among several colony characters in the honey bee (Hymenoptera: Apidae) taking queen and worker effects into account. Ann Entomol Soc Am 84:324–331

Bienefeld K, Reinhardt F, Pirchner F (1989) Inbreeding effects of queen and workers on colony traits in the honey bee. Apidologie 20:439–450

Bijma P (2011) A general definition of the heritable variation that determines the potential of a population to respond to selection. Genetics 189:1347–1359

Bijma P, Rutten M (2002) Lecture notes for the SelAction workshop. Wageningen University, Wageningen

Brascamp EW, Bijma P (2014) Methods to estimate breeding values in honey bees. Genet Sel Evol 46:53

Brascamp EW, Bijma P (2019) A note on genetic parameters and accuracy of estimated breeding values in honey bees. Genet Sel Evol 51:71

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

Büchler R, Andonov S, Bienefeld K, Costa C, Hatjina F, Kezic N et al. (2013) Standard methods for rearing and selection of *Apis mellifera* queens. J Apicult Res 52:1–30

Bulmer M (1971) The effect of selection on genetic variability. Am Nat 105:201–211

Chapman RE, Bourke AFG (2001) The influence of sociality on the conservation biology of social insects. Ecol Lett 4:650–662

Chevalet C, Cornuet J-M (1982) Étude théorique sur la sélection du charactère production de miel chez l'abeille. Apidologie 13:39–65

Cobey SW, Tarpy DR, Woyke J (2013) Standard methods for instrumental insemination of *Apis mellifera* queens. J Apicult Res 52:1–18

Dekkers JCM (1992) Asymptotic response to selection on best linear unbiased predictors of breeding values. Anim Prod 54:351–360

Dempfle L (1974) A note on increasing the limit of selection through selection within families. Genet Res 24:127–135

Fisher RA (1918) The correlations between relatives on the supposition of Mendelian inheritance. Trans R Soc Edinb 52:321–341

Gerula D, Węgrzynowicz P, Panasiuk B, Bieńkowska M, Skowronek W (2014) Performance of bee colonies headed by queens instrumentally inseminated with semen of drones who come from a single colony or many colonies. J Apicult Sci 58:87–97

Gianola D, Fernando RL, Im S, Foulley JL (1989) Likelihood estimation of quantitative genetic parameters when selection occurs: models and problems. Genome 31:768–777

Harbo, JR (1999) The value of single-drone inseminations in selective breeding of honey bees. In: Connor L, Hoopingarner R (eds) Apiculture for the 21st Century. Wicwas Press, Cheshire. pp 1–5.

Hawthorne D (1998) Predicting insect adaptation to a resistant crop. J Econ Entomol 91:565–571

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 31:423–447

Jensen JLWV (1906) Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Math 30:175–193

Kenney JF, Keeping ES (1957) Mathematics of statistics. Van Nostrand, New York

Le Conte Y, Navajas M (2008) Climate change: impact on honey bee populations and diseases. Rev Sci Tech OIE 27:499–510

Mikheyev AS, Tin MMY, Arora J, Seeley TD (2015) Museum samples reveal rapid evolution by wild honey bees exposed to a novel parasite. Nat Commun 6:7991

Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH (2002) BLUPF90 and related programs (BGF90). In: Proceedings of 7th world congress on genetics applied to livestock production, 28, pp. 743–744

Moritz RFA (1986) Comparison of within-family and mass selection in honeybee populations. J Apicult Res 25:146–153

Muñoz I, Dall'Olio R, Lodesani M, De La Rúa P (2014) Estimating introgression in Apis mellifera siciliana populations: are the conservation islands really effective? Insect Conserv Divers 7:563–571

Oldroyd BP, Rinderer TE, Harbo JR, Buco SM (1992) Effects of intracolonial genetic diversity on honey bee (Hymenoptera: Apidae) colony performance. Ann Entomol Soc Am 85:335–343

Pearson K (1903) Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. Philos Trans R Soc Lond A 200:321–330

Pernal SF, Sewalem A, Melathopoulos AP (2012) Breeding for hygienic behaviour in honeybees (*Apis mellifera*) using free-mated nucleus colonies. Apidologie 43:403–416

Pieplow JT, Brauβe J, van Praagh JP, Moritz RFA, Erler S (2017) A scientific note on using large mixed sperm samples in instrumental insemination of honeybee queens. Apidologie 48:716–718

Plate M, Bernstein R, Hoppe A, Bienefeld K (2019a) Comparison of infinitesimal and finite locus models for long-term breeding simulations with direct and maternal effects at the example of honeybees. PLoS ONE 14:e0213270

Plate M, Bernstein R, Hoppe A, Bienefeld K (2019b) The importance of controlled mating in honeybee breeding. Genet Sel Evol 51:74

Plate M, Bernstein R, Hoppe A, Bienefeld K (2020) Long-term evaluation of breeding scheme alternatives for endangered honeybee subspecies. Insects 11:404

Roehe R, Kennedy BW (1993) Effect of selection for maternal and direct genetic effects on genetic improvement of litter size in swine. J Anim Sci 71:2891–2904

Spivak M, Reuter GS (2001) Varroa destructor infestation in untreated honey bee (Hymenoptera: Apidae) colonies selected for hygienic behavior. J Econ Entomol 94:326–331

Tallis GM, Leppard P (1987) The joint effects of selection and assortative mating on a single polygenic character. Theor Appl Genet 75:41–45

Tallis GM, Leppard P (1988) The joint effects of selection and assortative mating on multiple polygenic characters. Theor Appl Genet 75:278–281

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

Turelli M, Barton NH (1990) Dynamics of polygenic characters under selection. Theor Popul Biol 38:1–57

Uzunov A, Brascamp EW, Büchler R (2017) The basic concept of honey bee breeding programs. Bee World 94:84–87

Uzunov, A, Büchler, R, Bienefeld, K (2015) Performance testing protocol. A guide for European honey bee breeders. Landesbetrieb Landwirtschaft Hessen, Kirchhain

van Boxtel C, van Heerden JH, Nordholt N, Schmidt P, Bruggeman FJ (2017) Taking chances and making mistakes: non-genetic phenotypic heterogeneity and its consequences for surviving in dynamic environments. J R Soc Interface 14:20170141

van der Werf JHJ, de Boer IJM (1990) Estimation of additive genetic variance when base populations are selected. J Anim Sci 68:3124–3132

Van Grevenhof EM, Van Arendonk JAM, Bijma P (2012) Response to genomic selection: the Bulmer effect and the potential of genomic selection when the number of phenotypic records is limiting. Genet Sel Evol 44:26

Wei M, Caballero A, Hill WG (1996) Selection response in finite populations. Genetics 144:1961–1974

Zayed A, Packer L (2005) Complementary sex determination substantially increases extinction proneness of haplodiploid populations. Proc Natl Acad Sci USA 102:10742–10746