Conservation of the dark bee (*Apis mellifera mellifera*): Estimating C-lineage introgression in Nordic breeding stocks


To link to this article: https://doi.org/10.1080/09064702.2020.1770327

Published online: 01 Jun 2020.

Article views: 102
Conservation of the dark bee (Apis mellifera mellifera): Estimating C-lineage introgression in Nordic breeding stocks


ABSTRACT
Displacement and admixture are threatening the survival and genetic integrity of the European dark bee, Apis mellifera mellifera. Studies on the phenotype-genotype map and genotype by environment interactions in honey bees are demonstrating that variation at subspecies level exists and is worth conserving. SNP-based tools for monitoring genetic integrity in bees have been developed, but are not yet widely used by European dark bee breeders. We used a panel of ancestry informative SNP markers to assess the level of admixture in Nordic dark bee breeding stocks. We found that bee breeders falsely classified admixed stocks based on morphometry as purebred and vice versa. Even though most Nordic A. m. mellifera breeding stocks have low proportions of C-lineage ancestry, we recommend to incorporate genotyping in Nordic dark bee breeding programmes to ensure that minimal genetic diversity is lost, while the genetic integrity of the subspecies is maintained.

ARTICLE HISTORY
Received 1 November 2019
Accepted 12 May 2020

KEYWORDS
Genetic diversity; beekeeping; brown bee; black bee; hybridisation; honey bee; admixture

Introduction
The honey bee, Apis mellifera, provides a wide range of benefits to humans, ranging from the production of honey, wax, propolis and royal jelly to the pollination of agricultural crops, with honey bees being the most important managed pollinator (Klein et al., 2007; Aizen & Harder, 2009; vanEngelsdorp & Meixner, 2010; Breeze et al., 2011; Potts et al., 2016).

Humans and bees have interacted for thousands of years and both natural and artificial selection have shaped contemporary bee populations (Ruttner, 1988; Roffet-Salque et al., 2015). The evolutionary history of the honey bee A. mellifera is not fully understood and there are hypotheses of an Asian as well as an African evolutionary origin, with the latest and most comprehensive study advocating an origin in the Middle East or Northern Africa (Whitfield et al., 2006; Han et al., 2012; Wallberg et al., 2014; Cridland et al., 2017). The species is found throughout Africa, Asia and Europe and subdivided into at least 27 morphologically, geographically, physiologically and behaviourally distinct subspecies (Ruttner, 1988; Parker et al., 2010; Dogantzis & Zayed, 2019), with new subspecies recently described (Sheppard & Meixner, 2003; Meixner et al., 2011; Chen et al., 2016). While there may not be a consensus in the scientific community on the origins and major migration events and routes of A. mellifera, it is broadly accepted that there are at least five genetically distinct evolutionary lineages: A lineage (Africa), Y lineage (Arabian Peninsula and Horn of Africa), O lineage (Asia) and the C and M lineage (Europe) (Dogantzis & Zayed, 2019). The European M lineage, traditionally comprised of A. mellifera ligustica (dark honey bee) and A. mellifera iberiensis, has its distribution from the Iberian peninsula in the South to southern Scandinavia in the North and from the British Isles in the West across much of Europe to the Ural mountains in the East (Ruttner, 1988). Recently, this distribution was extended eastwards by the description of a new M-lineage subspecies, A. mellifera sinisxinyuan, from China (Chen et al., 2016). The C lineage from the Central Mediterranean and Southeastern European region is more restricted in its range, even though its subspecies are more numerous, with the two apiculturally most important subspecies A. mellifera ligustica and A. mellifera carnica found on the Italian and the Balkan peninsula, respectively (Ruttner, 1988).
Due to human apicultural activity, the distribution and the genetic integrity of many subspecies no longer resemble the state described above. In Europe, across large geographic areas, M-lineage bees have been replaced by C-lineage bees, mostly *A. m. carnica* and *A. m. ligustica*, and the remaining M-lineage bees are threatened due to high levels of introgression (Jensen et al., 2005; Soland-Reckeweg et al., 2009; Oleksa et al., 2011; Chávez-Galarza et al., 2013; Pinto et al., 2014; Muñoz et al., 2015; Parejo et al., 2016).

Admixture in the honey bee has an impact on phenotypic traits and adaptation. The poster child example for this are so-called Africanised bees, African A-lineage bees which hybridised with originally European M- and C-lineage bees throughout South – and later North America (Smith et al., 1989; Rinderer et al., 1991; Sheppard et al., 1991; Pinto et al., 2005). These bees have caused many problems for beekeepers, as they are difficult to manage and more defensive than the bees of European origin (Michener, 1975; Rinderer, 1986; Winston, 1992; Ferreira et al., 2012).

The number of studies aiming to shed light on the genotype-phenotype map are increasing. Recent studies, employing QTL mapping, GWAS and genome-wide scans, have focused on pathogen resistance (Behrens et al., 2011; Tsuruda et al., 2012; Behrens & Moritz, 2014; Huang et al., 2014; Spötter et al., 2016), colony defence (Hunt et al., 1998; Guzmán-Novoa et al., 2002; Avalos et al., 2017), age at first foraging (Rueppell, 2009), high-altitude adaptation (Wallberg et al., 2017), behavioural specialists within a colony (Southey et al., 2016), royal jelly production (Wragg et al., 2016) and the detection of selection signatures in general (Zayed & Whitfield, 2008; Chávez-Galarza et al., 2013; Parejo et al., 2017; Henriques, et al., 2018b). Other studies have concentrated on elucidating genotype by environment interactions (Costa et al., 2012a; Meixner et al., 2014; Francis et al., 2014b). Significant interactions between the genetic origin and the environment were found for colony survival (Büchler et al., 2014), pathogen levels (Francis et al., 2014a), colony development (Hatjina et al., 2014), spring development and honey production (Costa et al., 2012b), whereas data for behavioural differences was not as clear cut (Uzunov et al., 2014).

From the assumption that *A. mellifera* subspecies differ genetically, physiologically and behaviourally, as suggested by the aforementioned studies, follows a need for conservation measures at the subspecies level, especially for *A. m. mellifera*, one of the most threatened subspecies. Following the publication of the honey bee genome and concurrent availability of powerful genomic resources, there have been efforts to develop single nucleotide polymorphism (SNP) tools for the detection of C-lineage introgression in M-lineage bees (Muñoz et al., 2015, 2017; Parejo et al., 2016; Henriques et al., 2018a; Henriques et al., 2019) and of African ancestry in European stocks in North America (Crozier et al., 1991; Chapman et al., 2015). These tools can be used by bee breeders to more accurately screen their breeding stocks and make breeding decisions accordingly.

In the Nordic countries, where *A. m. mellifera* is native, the majority of honey bee colonies today are *A. m. carnica*, *A. m. ligustica*, or the intended hybrid ‘Buckfast’. There are legally protected conservation areas in Norway and Denmark and all four countries, Denmark, Finland, Sweden and Norway, have dark bee associations dedicated to the conservation of the subspecies. The number of *A. m. mellifera* colonies in Finland and Denmark is very low, with <300 colonies per country, whereas the populations in Sweden and Norway consist of a few thousand colonies (Ruottinen et al., 2014). The Nordic Genetic Resource Center supports the ongoing conservation work by maintaining a network of Nordic dark bee breeders (Demant et al., 2019).

Currently, dark bee breeders in the Nordic countries are largely relying on the combination of a few indexes based on wing venation, such as cubital index, Hantel index and discoidal shift angle, as well as habits and presumed ancestry for making breeding decisions (Bouga et al., 2011). Whether the use of these traits has been effective in preserving the genetic integrity of *A. m. mellifera* across the Nordic countries is unclear. To address this issue, in this study we inferred admixture proportions in the Nordic dark bee breeding stocks, using one of the available SNP-based tools for estimating C-lineage introgression in M-lineage bees (Henriques et al., 2018a), and compared the results with the dark bee breeders’ own assessment. Furthermore, we were interested in contrasting the two types of data, genotype and wing morphometric data, at the colony level by analysing multiple drones per colony. At the same time, we wanted to collect morphometric data from purebred drones, identified as such by the genotype data, for future use as reference material. If possible, we also wanted to compare our morphometric drone data with any available public drone data, in order to corroborate thresholds for purity and compare our morphometric with genotypic data.

**Materials and methods**

**Sampling**

Both female (worker) bees and male (drones) were collected for this study (Table 1). Genetic analyses were...
carried out on both the worker and the drone samples, whereas morphometric measurements were only taken from drone samples. Twenty newly hatched worker bee individuals per colony from 19 colonies were collected in a total of four apiaries in Norway, Sweden and Finland. Ten newly hatched drone individuals per

### Table 1. Drone and worker bee samples used in this study.

<table>
<thead>
<tr>
<th>Colony</th>
<th>Lon</th>
<th>Lat</th>
<th>Caste</th>
<th>Breeder classification</th>
<th>#indiv sampled</th>
<th>#indiv morpho-metrics</th>
<th>#indiv genotypes</th>
<th>Classification genotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>LK1</td>
<td>10.4974</td>
<td>59.7768</td>
<td>drone</td>
<td>hybrid</td>
<td>10</td>
<td>9</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>LK2</td>
<td>10.4974</td>
<td>59.7768</td>
<td>drone</td>
<td>purebred</td>
<td>10</td>
<td>9</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>LK3</td>
<td>10.4974</td>
<td>59.7768</td>
<td>drone</td>
<td>hybrid</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>LK4</td>
<td>10.4974</td>
<td>59.7768</td>
<td>drone</td>
<td>hybrid</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>LK5</td>
<td>10.4974</td>
<td>59.7768</td>
<td>drone</td>
<td>hybrid</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>LK6</td>
<td>10.4974</td>
<td>59.7768</td>
<td>drone</td>
<td>hybrid</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>LK7</td>
<td>10.4974</td>
<td>59.7768</td>
<td>drone</td>
<td>hybrid</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>LK8</td>
<td>10.4974</td>
<td>59.7768</td>
<td>drone</td>
<td>hybrid</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>PR1</td>
<td>8.5074</td>
<td>59.1765</td>
<td>drone</td>
<td>purebred</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>PR2</td>
<td>8.5074</td>
<td>59.1765</td>
<td>drone</td>
<td>purebred</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>PR3</td>
<td>8.5074</td>
<td>59.1765</td>
<td>drone</td>
<td>purebred</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>PR4</td>
<td>8.5074</td>
<td>59.1765</td>
<td>drone</td>
<td>purebred</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>PR5</td>
<td>8.5074</td>
<td>59.1765</td>
<td>drone</td>
<td>purebred</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>PR6</td>
<td>8.5074</td>
<td>59.1765</td>
<td>drone</td>
<td>purebred</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>PR7</td>
<td>8.5074</td>
<td>59.1765</td>
<td>drone</td>
<td>purebred</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

* denotes where the breeder’s classification does not correspond to the classification based on genotypes.
colony from 48 colonies were collected from seven apiaries (one of which is also represented in the worker bee sample) in Norway (Figure 1). All samples were preserved in 100% acetone (drones) or 96% ethanol (worker bees). All of the participating bee breeders are interested in conserving the Nordic dark bee, and presumably have purebred *A. m. mellifera* colonies. They are all members of their respective national dark bee beekeeping associations, which support their members in using wing morphometry for breeding decisions. The bee breeders classified their colonies as purebred or hybrid a priori based on presumed ancestry, habitus and to some extent on wing indexes, such as cubital index, Hantel index and discoidal shift angle.

**Reference data**

**Morphometric reference data**

Morphometric reference data for drones was only available for one of our measurements, the cubital index (Table S1). Data from 28 *A. m. carnica*, 4 *A. m. ligustica* and 11 *A. m. mellifera* colonies, consisting of means of at least 10 individuals, were obtained from the Institut für Bienenkunde, Oberursel, Germany. This institute houses the largest compilation of reference data for morphometric characters for the majority of honey bee subspecies based on the work of Ruttner (1988) and Meixner et al. (2013).

**Genotype reference data**

Genotype reference data for C- and M-lineage bees were taken from Henriques et al. (2018a), supplementary material. Samples were chosen to maximise geographic spread and minimise introgression and missing data. Eight drone samples of *A. m. ligustica*, *A. m. carnica* and *A. m. mellifera*, as well as 18 and 20 worker bee samples of *A. m. carnica* and *A. m. mellifera*, respectively, were compiled as reference data sets for the analyses carried out in this study (Table S2).

**Morphometric measurements**

Morphometric analyses were performed using the right forewings of 463 drone samples from 48 colonies from seven apiaries (Table 1). The wings were mounted on slides (Gepe Geimuplast GmbH) and scanned on a flatbed scanner (Epson Perfection V600 Photo) with a resolution of 2400 ppi. The scanned images were analysed.
using the software DrawWing (Tofilski, 2004) and the cubital index (Ci), discoidal shift angle (DsA) and Hantel index (Hi) were determined, as well as any vein anomalies noted (Smith et al., 1997). All statistical tests on wing morphometric data were carried out in R version 3.5.2 (R Core Team, 2018) employing the inbuilt stats base package (one-way ANOVA, Tukey HSD test, χ2) and the MASS_7.3-51.1 package (Fisher’s Linear Discriminant Analysis).

**DNA extraction and genotyping**

DNA was extracted from 380 workers and 375 drone samples. Briefly, thorax tissue was excised and dried in RT, followed by a 2 h incubation at 55°C with 400 μg Proteinase K (LGC Genomics) and 4 μg Ambion™ RNaseA (Thermo Fisher Scientific #AM2270). DNA was extracted from the lysate according to the manufacturer’s instructions (sbeadex livestock kit #44701/44702, LGC Genomics) and a MagMax Express 96 (Life Technologies).

Samples with sufficient yields were normalised to 15 ng/μL in a total volume of 30 μL and sent for genotyping to the Genomics Unit of the Instituto Gulbenkian de Ciência, Portugal. Two assays (M1 + M3; Henriques et al., 2018a) totalling 62 SNPs were run using the iPLEX chemistry on a MassARRAY® MALDI-TOF (Agena BioScience 2018a) totalling 62 SNPs were run using the iPLEX chemistry on a MassARRAY® MALDI-TOF (Agena BioScience 2018a) for drones. Brieﬂy, thorax tissue was excised and dried in RT, followed by a 2 h incubation at 55°C with 400 μg Proteinase K (LGC Genomics) and 4 μg Ambion™ RNaseA (Thermo Fisher Scientific #AM2270). DNA was extracted from the lysate according to the manufacturer’s instructions (sbeadex livestock kit #44701/44702, LGC Genomics) and a MagMax Express 96 (Life Technologies).

Samples with sufficient yields were normalised to 15 ng/μL in a total volume of 30 μL and sent for genotyping to the Genomics Unit of the Instituto Gulbenkian de Ciência, Portugal. Two assays (M1 + M3; Henriques et al., 2018a) totalling 62 SNPs were run using the iPLEX chemistry on a MassARRAY® MALDI-TOF (Agena BioScience 2018a) for drones. Brieﬂy, thorax tissue was excised and dried in RT, followed by a 2 h incubation at 55°C with 400 μg Proteinase K (LGC Genomics) and 4 μg Ambion™ RNaseA (Thermo Fisher Scientific #AM2270). DNA was extracted from the lysate according to the manufacturer’s instructions (sbeadex livestock kit #44701/44702, LGC Genomics) and a MagMax Express 96 (Life Technologies).

**SNP quality control and filtering**

PLINK v1.90b5 (Chang et al., 2015) was used for filtering and merging the raw SNP (Table S3) and reference data. The combined reference and raw data sets, for drones and workers separately, were filtered for missing data, with SNPs and samples with missing call rates above 10% being removed.

**Admixture estimation**

The software ADMIXTURE v 1.3 (Alexander et al., 2009) was used to estimate ancestry proportions (Q-values) for K = 2 for the worker and drone (haploid= ‘male:*’) data set separately. The ancestry informative marker panel used was designed for distinguishing between M- and C-lineage ancestry (Muñoz et al., 2015; Henriques et al., 2018a), such that only K = 2 was chosen. Fifty replicates, with a random seed generated from current time, were run unsupervised with the default optimisation method, a block relaxation algorithm, and the default termination criterion of stopping when the log-likelihood increased by less than ε = 10^-4 between iterations. CLUMPAK was used to summarise and visualise the obtained ancestry estimates using default settings (Kopelman et al., 2015). We followed Henriques et al. (2018a) in defining purebred A. m. mellifera individuals as those with a Q-value < 0.05 for C-lineage ancestry. Any individual with a Q-value between >0.05 and <0.95 for C-lineage ancestry was considered a hybrid and any colony containing one or more hybrid individual was scored as ‘hybrid’.

**Results**

**Morphometrics – drones**

Inspection of the cubital index reference data showed that the three subspecies A. m. ligustica, A. m. carnica and A. m. mellifera do not represent discrete groups in this character (Figure 2(a)). The same is true, when comparing C-lineage bees (A. m. ligustica + A. m. carnica) with M-lineage bees (A. m. mellifera), which due to low sample numbers for A. m. ligustica may be sensible (not shown).

The means of the three groups differ significantly at the p < .001 level according to a one-way ANOVA [F (2,40) = 28.34, p = 2.16e−08]. Post hoc comparisons using the Tukey HSD test indicated that the mean cubital index values of A. m. mellifera (M = 1.48, SD = 0.18) was significantly different from both A. m. ligustica (M = 1.77, SD = 0.247; p = .019) and A. m. carnica (M = 1.95, SD = 0.167; p = .000) but that the latter two did not differ from one another. Nonetheless, it was not possible to reliably predict group membership based on the cubital index of drones alone. A Fisher’s Linear Discriminant Analysis with the reference samples yielded a 16.3% and 9.3% misclassification rate for the three subspecies and two lineages, respectively.

When plotting colony averages for C-lineage ancestry proportions (Q-values) against colony averages for the cubital index for M- and C-lineage reference samples as well as our samples, it became clear that the cubital index measured on drones is not reliable in detecting hybrid individuals, since a number of colonies with high average C-lineage ancestries between 0.09 and 0.32 (n = 5) have cubital index values that are well within the range (<3rd quantile) of the reference M-lineage values (Figure 3).

This demonstrated that the available public data, of characters commonly measured by Nordic dark bee breeders, is not sufficient for establishing thresholds, which could then be used for separating purebreds from hybrids based on morphometric data.

If we accept the genotype data as the decisive factor for subspecies determination, our morphometric data
can be used as reference data in the future. Our cubital index data from purebred colonies does not significantly differ from the reference data according to a one-way ANOVA \( F(1,33) = 0.006, p = .941 \); Figure 2(b). Morphometric measurements (Ci, DsA, Hi and vein anomalies) of all individuals from colonies classified as purebred according to genotype data are given in Table S4a + b.

In our sample of 405 individuals from 42 colonies, for which both morphometric and genotype data were available, a total of 24 vein anomalies were noted. Of these 24, 18 were identified in 10 hybrid colonies (at least one admixed individual) and six in three purebred A. m. mellifera colonies (Table 2). Individuals from hybrid colonies are more likely to exhibit anomalies than those from purebred colonies based on a comparison of ‘no anomalies’ vs. ‘at least one anomaly per colony’ \( \chi^2 (1, N = 42) = 7.02, p < .01 \).

Admixture estimation

No genotypes were obtained for six colonies (LK1-LK6) due to human error. In the filtering process, 23 individuals and 1 SNP (AMB-00524451) were removed for the

<table>
<thead>
<tr>
<th>Table 2. Number of wing venation anomalies for hybrid and purebred A. m. mellifera colonies based on measurements of drone wings.</th>
</tr>
</thead>
<tbody>
<tr>
<td>hybrid</td>
</tr>
<tr>
<td># colonies</td>
</tr>
<tr>
<td># individuals</td>
</tr>
<tr>
<td># anomalies</td>
</tr>
<tr>
<td># colonies no anomaly</td>
</tr>
<tr>
<td># colonies at least one anomaly</td>
</tr>
</tbody>
</table>

Note: Colonies with at least one individual with a Q-value between >0.05 and <0.95 for C-lineage ancestry were considered hybrid.
drone data set, whereas 3 individuals were removed for the worker data set, leaving a drone data set with 327 individuals and 61 SNPs and a worker data set with 415 individuals and 62 SNPs (Table 3).

No minor modes were identified for 50 ADMIXTURE runs of either the drone or worker data set, meaning that only one genuine solution was found for both data sets. These solutions are depicted in Figure 4 for workers (a) and drones (b) separately.

According to the bee breeders, the worker samples from 19 colonies were all deemed to be purebred A. m. mellifera. We found evidence of C-lineage ancestry in five of the 19 colonies (Table 1), with the number of introgressed individuals per colony ranging from one (KD2, Q-value: 0.05) to all 20 (AN13, Q-values: 0.07–0.27; AN14, Q-values: 0.05–0.16). Both the number of detected introgressed individuals as well as the proportion of C-lineage ancestry varied, but there was no evidence for F1 hybrids, which would be expected to have Q-values around 0.50 (Henriques et al., 2018a).

Of the 42 colonies from which we sampled drones and obtained genotype data for, 37 were classified by the bee breeders as purebred and five as hybrid colonies. Our analyses confirmed the bee breeders’ classification of 25 colonies; of which three were hybrid and 22 purebred A. m. mellifera (Table 1). A total of 17 colonies were not found to be in agreement with the bee breeders’ classification. We found evidence of C-lineage ancestry in 15 colonies that were deemed purebred A. m. mellifera by the bee breeders, with the number of hybrid individuals per colony ranging from two out of eight (SS1, Q-values: 0.05 and 0.06) to seven out of seven (TR4, Q-values: 0.07–0.22). Average C-lineage Q-values per colony ranged from 0.03 (SS1) to 0.32 (TR2), with seven of the 15 colonies below the 0.05 threshold. Furthermore, two colonies turned out to be purebred, which were thought to be hybrid prior to our analyses.

Across data sets, 38 colonies were found to be purebreds and 23 hybrids. Of the hybrid colonies, 15 generally demonstrated low proportions of C-lineage ancestry (Q-values < 0.13) and often only for a few individuals within the colony. In contrast, the colonies LK7 and LK8, classified as hybrid already by the bee breeder and AN13, AN14, TR1, TR2, TR3 and TR4 were found to have higher proportions (Q-values ≥ 0.13) of C-lineage ancestry. For example, of the seven TR2 samples, five had Q-values between 0.42 and 0.57, strongly suggesting that they represent F1 hybrids (Figure 4(a,b)).

---

**Table 3.** Number of colonies, number of individuals with and without reference individuals, number of individuals/SNPs removed during filtering (10% missing data) and number of SNPs in the final data sets.

<table>
<thead>
<tr>
<th>Final data sets</th>
<th># colonies</th>
<th># individuals included (with/without reference)</th>
<th># individuals/SNPs removed</th>
<th>Reference subspecies</th>
<th># SNPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drones</td>
<td>42</td>
<td>327/304</td>
<td>23/1</td>
<td>carnica, ligustica, mellifera</td>
<td>61</td>
</tr>
<tr>
<td>Workers</td>
<td>19</td>
<td>415/378</td>
<td>3/0</td>
<td>carnica, mellifera</td>
<td>62</td>
</tr>
</tbody>
</table>

---

**Figure 4.** Individual ancestry estimates for worker samples from 19 colonies (a) and for drone samples from 42 colonies (b) for $K = 2$. Mean similarity score and mean LogLikelihood value of the 50 replicate runs is given. Each bar represents an individual and colonies, labelled at the bottom, are separated by black lines. For colonies marked with an asterisk at least one individual was found to be a hybrid. Colours correspond to the two presumed ancestries, with M-lineage ancestry in blue and C-lineage ancestry in orange.
Discussion

Reliable reference data are crucial for using morphometric data to determine subspecies status. Traditionally, diploid worker bees have been used to ascertain subspecies membership based on morphometry (Bouga et al., 2011) and there is thus very little reference data available for drones. Haploid drones are generally less suitable for this purpose because they exhibit greater variability in morphological characteristics, such as wing venation, and higher numbers of vein anomalies than workers (Casteel & Phillips, 1903). However, in breeding practice, there are situations where the direct analysis of drones, thus permitting a direct judgment of the queen, is helpful; for example, choosing drone mother colonies. When choosing mother colonies for supplying drones to isolated dark bee mating stations, whether the queen has mated with one or more drones from another subspecies is of no importance. If one uses worker bees to assess the purity of the drone mother colony, there is a chance that a purebred queen is discarded due to crosses with one or more non-purebred drones. When breeding in small populations for conservation purposes, one does not want to discard valuable genetic diversity unnecessarily, in which case drones become the most suitable tool to assess the purity of the mother queen. We could not find any public reference data for drones for the discoidal shift angle and Hantel index and the sample size per subspecies for the cubital index was very small, thus preventing us from defining subspecific thresholds and subsequently directly comparing morphometric with genotypic data of our drone sample. It has been shown that geometric morphometrics of workers and drones in A. m. iberiensis are largely congruent and provide similar information on the genetic structure as assessed by SNP data (Henriques et al., 2020). This type of study, either carried out on drones with many more morphometric characters and one’s own drone reference samples or on worker bees, would be very valuable in our eyes. We propose to accept our morphometric drone data from purebred A. m. mellifera, as defined by genotype data, as reference data for the subspecies.

Genotype data has the power to reveal admixture events that happened many generations ago, but precisely timing hybridisation events on an evolutionary timescale is difficult (Strasburg & Rieseberg, 2011; Payseur & Rieseberg, 2016). Contemporary hybridisation events, such as F1 hybrids of C- and M-lineage bees, have been experimentally shown to have Q-values ≈ 0.5, as expected from theory (Henriques et al., 2018a). Our results showed that there were misclassifications by the bee breeders in both directions: colonies deemed purebred A. m. mellifera were found to contain admixed individuals, and supposed hybrid colonies turned out to be purebred, according to the ancestry informative markers employed. Here one has to keep in mind that our classification was rather conservative, with labelling a colony as hybrid when just one individual within the colony was shown to have a C-lineage ancestry proportion over 0.05. The queen of a colony, where only a few worker individuals of a sample were found to be admixed, can of course be pure, with the C-lineage ancestry in the admixed worker bees (offspring) coming from the drone side (paternal side). Nonetheless, the misclassifications by the bee breeders may lead to valuable genetic diversity being discarded, which could be used for breeding, because it is falsely assumed that they are admixed stocks. This is especially detrimental in a species where a loss of diversity in sex alleles can lead directly to decreased brood viability (Woyke, 1980, 1981; Hyink et al., 2013). On the other hand, we identified colonies that bee breeders would use in breeding, which contained admixed individuals, including F1 hybrids, which the bee breeder was unaware of. This is a problem when aiming to maintain the genetic integrity of the subspecies. Moreover, if one would like to market dark bee products, there should be a reliable way to monitor that the claims being made are in fact correct.

One possible pitfall in sampling and genotyping bees is that individuals from other colonies may enter a hive and thus be sampled. This is especially true for drones (Moritz & Neumann, 1996). We hope that this error is minimised in our sample due to sampling newly hatched drones; but we cannot fully exclude this possibility.

So far, bee breeders have had to rely on their own judgements and morphometric measurements when selecting breeding material. Many breeders are not aware that they may be discarding purebred queens with good behavioural qualities, along with valuable genetic diversity. A direct comparison of morphometric with genotypic data could exemplify this.

In order to make SNP-based tools attractive for monitoring the genetic integrity of breeding stocks, there are a number of practical issues that need to be resolved. The turnaround time of sampling until receiving understandable results needs to be aligned with the beekeeping year. When breeding is based on performance testing roughly a year is available, since queens will be tested and sampled in year 1 and breeding based on the results will be carried out in year 2. Costs need to be low enough that this investment pays off. And most of all, the sampling and genotyping procedure need to be simple and well described and the results understandable to a layman.
The Nordic countries are party to the 1992 United Nations Convention on Biological Diversity and the Food and Agriculture Organization of the United Nations’ (FAO) Global Plan of Action for Animal Genetic Resources (FAO, 2007) and thus are committed to the ‘conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of benefits arising from genetic resources’ (United Nations, 1992). This includes both wild and managed genetic resources. Even though most A. mellifera colonies are managed, they have largely not been classified as farm animals in the past. This may currently be changing. The FAO considers honey bees to be animal genetic resources for food and agriculture, similar to livestock when it comes to genetic diversity and breeding. Efforts are underway to add the honey bee to FAO’s worldwide Domestic Animal Diversity Information System (DAD-IS), which includes a searchable database containing information such as the status of breeds regarding their risk of extinction (Food and Agriculture Organization of the United Nations, 2019, para. 92). We hope that this new classification may clarify responsibilities at the national level and make new funding sources available for conservation efforts. Extra funding could permit the Nordic dark bee breeding programmes to complement their currently standard morphometric work with genotype data, thus ensuring that the least amount of genetic diversity is discarded and the least amount of admixed stock is used in breeding.

**Acknowledgements**

We are greatly indebted to the participating bee breeders for allowing us to include their samples in this study and for providing the samples. We sincerely thank Dr. Stefan Fuchs for providing the morphological reference data. This project would not have been possible without the help of João Costa from the Genomics Unit of the Instituto Gulbenkian de Ciência, Portugal, who carried out the genotyping for us. We thank the two anonymous reviewers who helped greatly to improve the manuscript. This research was partly funded by the Norwegian Agriculture Agency 17/3472.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was supported by the Norwegian Agriculture Agency [grant number 17/3472].

**Data availability statement**

All data is available in the supplementary material.

**References**


