Apidologie (2023) 54:24

https://doi.org/10.1007/s13592-023-00999-w

**Original article**

# Founder gene pool composition and genealogical structure in two populations of Austrian Carniolan honey bees (*Apis mellifera carnica*) as derived from pedigree analysis

Thomas Druml[1], Anselm Putz[2], Michael Rubinigg[3], Martin H. Kärcher[4], Karl Neubauer[5], and Christian Boigenzahn[3]

[1] BBS, Marktstraße 8 GS01, 9601 Arnoldstein, Austria
[2] Auerbach 34, 8242 St. Lorenzen am Wechsel, Austria
[3] Biene Österreich, Georg Couch Platz 1, 1010 Wien, Austria
[4] Austrian Carnica Association (ACA), Leonhardstraße 114/35, 8010 Graz, Austria
[5] Zentrale Arbeitsgemeinschaft der Carnicazüchter (ZAC!), Pachmayrstraße 57, 4040 Linz, Austria

Received 3 November 2022 – Revised 9 February 2023 – Accepted 21 February 2023

**Abstract** – Pedigree analyses describing gene pool and genetic diversity frequently have been performed for multiple livestock species. In honey bees, comparable studies are not yet available, and therefore, we aimed to investigate the genetic diversity of two Austrian *Apis mellifera carnica* breeding populations by means of pedigree analysis. Honey bee breeding programs in Austria get administered by two breeding associations, the ACA and the ZAC!. Their respective reference populations comprised the birth years 2019 and 2020 and resulted in 2.675 breeding queens within the ACA and 1.286 queens within the ZAC! population. From the total of 1.015 ACA founder queens, 13 founders represented 50% of the gene pool; for the ZAC! population (624 founders in total), 21 founders were responsible for 50% of the segregating alleles. The genetic diversity indices like effective numbers of founders ($f_e$), ancestors ($f_a$), and founder genome equivalents ($n_g$) are capable to determine unbalanced breeding practices, occurrence of genetic bottlenecks, and genetic drift in the respective population histories. The values obtained (ACA/ZAC!: $f_e = 71/125$; $f_a = 30/48$; $n_g = 18.7/21.6$) demonstrated genetic loss due to unbalanced, excessive use of single breeding animals within the ACA population. The slightly lower loss in diversity within the ZAC! population can be attributed to a smaller active population number. As a consequence, both populations exhibit moderate decrease of genetic diversity, which is comparable to mammal livestock with small or limited population size.

***Apis mellifera carnica* / Genetic diversity / Gene pool / Founder contribution / Pedigree analysis**

## 1. INTRODUCTION

The Carniolan honey bee (*Apis mellifera carnica*), shortly named Carnica, can be considered as the predominantly used genetic stock for queen breeding and honey production in Central, Eastern, and Southeastern Europe (Lodesani and Costa 2003; Hoppe et al. 2020). The area of origin of this subspecies is defined by the natural borders of the alps in the northwest, the Carpathian Mountains in the northeast, the Mediterranean Sea in the southwest, and the Albanian alps in the south (Ruttner 1988). From an evolutionary and phylogenetic point of view, *A. m. carnica* belongs to the so-called C-lineage, a genetic cluster which developed about the last

Corresponding author: T. Druml, thomas.druml@gmx.at
Handling Editor:

glacial maximum (13.000–38.000 BC) out of the common Middle Eastern ancestor *A. mellifera* (Wallberg et al. 2014). Beside natural selection, humans have contributed to shaping the current diversity of honey bees by means of migration, novel management techniques, and selective breeding.

In Austria, selective breeding on a broader scale, i.e., with significant impact on population structure, started between the beginning and middle of the twentieth century. Here, single breeders established local founder populations by mass selection within the first generations and line breeding programs in the following generations. These breeding programs were all based upon the establishment of so-called mating stations. Further on, the discovery of multiple mating in the honey bee in 1954 (Ruttner 1956) led to the use of multiple drone-producing daughter queens descending from one sire colony (4a) per mating station and to the establishment of pedigree nomenclature and studbook keeping in honey bee breeding.

After World War II, Prof. Friedrich Ruttner introduced breeding and performance testing programs following modern principles of animal breeding and genetics (Ruttner 1972). In order to control inbreeding growth, the mating concepts were based on rotating line breeding or breeding in semi-isolated groups (Pirchner cited in: Ruttner 1996). In the early 1990s, Prof. Kaspar Bienefeld introduced BLUP breeding value estimation (EBV) methods in Germany and substantially initiated a change in contemporary international bee breeding concepts (Bienefeld et al. 2007). Parallel to the implementation of EBV in Germany from 1992 to 1994, the database system BeeBreed (www.beebreed.eu) has been developed, which subsequently advanced to an online database tool enabling breeders and breeding associations to store and manage pedigree and phenotypic data necessary for the yearly calculations of breeding values. Since 2002, the Institute for Bee Research Hohen Neuendorf e.V. (LIB) also administrates the breeding programs for foreign breeding associations and several honey bee breeds and populations, respectively (Hoppe et al. 2020). Thus, the basic concepts of contemporary international EBV-assisted breeding programs rely on mass selection within semi-open studbooks and on the maximization of breeding values by simultaneous control of inbreeding growth (Bienefeld et al. 2007; Hoppe et al. 2020).

In Austria, Carnica breeders joined together in 1992 into the ACA breeding association (Austrian Carnica Association), an association that was in the beginning strongly associated to the Honey Bee Research Center Lunz am See. A breeding program including EBV methods was established in 1994 following a method developed by Willam and Essl (1993) using selection index theory according to Lush (1947). In the years 2002 and 2003, the ACA split into two breeder associations—the Zentrale Arbeitsgemeinschaft Österreichischer Carnicazüchter (ZAC!) and the ACA. In 2002, the ACA joined the BeeBreed administered international breeding programs. The ZAC! further developed the original breeding program of the ACA and switched to EBV following methods of Brascamp and Bijma (2014) and Brascamp et al. (2016) since 2017. Pedigree records and phenotype information of the ZAC! are managed by the online database BeeData (www.biene-oesterreich.at), established in 2002. As a consequence, two separate populations of *A. mellifera carnica* developed in Austria, both mainly originating from a same founder population existing between 1992 and 2002. In Austria, 390.607 honey bee colonies (including estimated number of 250.000 colonies of *A. m. carnica*) have been registered in total for the year 2019 (Boigenzahn et al., 2020). Within the same year, 1.529 colonies took part in the breeding program of the ACA and 452 colonies were tested in the ZAC! breeding program. As a result, about 1.0% of Austrian Carnica honey bee population is involved in controlled breeding. Contemporary Austrian honey bee mating systems are firmly founded on controlled and legally certified mating stations, mostly located in isolated alpine valleys. For the ACA, in total, 32 mating stations (28 to 30 stations are considered "purebred" stations, i.e., using several full sib daughters out of one 4a colony) were used for 2019/2020, and the ZAC! administrated 10 alpine "purebred" mating stations throughout 2019/2020. From, in total, 2.675 ACA queens born in 2019 and 2020, 2007 queens (75%) were mated

on controlled mating stations. 337 queens (12.6%) got inseminated, and 290 ACA queens (10.8%) were mated in an uncontrolled way. Within the ZAC! breeding program, artificial insemination plays a minor role, and more than 90% of 1.286 queens were mated on controlled "purbred" mating stations.

Pedigree information and its analysis represent a special field of animal breeding science. Throughout the last three decades, numerous publications have described the genetic structure and the effects of selection programs of different livestock breeds based on the analysis of levels of inbreeding and founder contributions (Cunningham 1991; Gandini et al. 1992; Moreaux et al. 1996; Boichard et al. 1997; Zechner et al. 2002; Aberle et al. 2003; Druml et al. 2009, 2016). Early (pre-digital) scientific pedigree analyses were limited to case studies of population samples and investigated the origins of British Shorthorn cattle by means of average inbreeding, coancestry within populations, coancestry of selected individuals to the population, coancestry between breeds or populations, and generation interval (Wright and McPhee 1925).

Another approach, based on the calculation of probabilities of gene origin, was proposed by Dickson and Lush (1933) for the first time and later on adapted for bigger pedigrees by McCluer et al. (1986). In the beginnings of the 1990s, Lacy (1989) introduced the parameters "effective number of founders" and "effective number of founder genomes" to describe genetic variability in zoo populations. A comprehensive survey of the state of the art in pedigree analysis was provided by Boichard et al. (1997), who also introduced the parameter "effective number of ancestors," thus accounting for bottlenecks in pedigrees.

In the meantime, the tools as described by Boichard et al. (1997) have frequently been applied to investigate gene pool, genetic diversity, inbreeding, and effects of selection programs in nearly all livestock species. Although pedigree documentation plays a major role in honey bee breeding, studies describing gene pool and genetic diversity of honey bee populations have not been published yet. In this study, we aim to investigate the genetic structure and genetic diversity of two Austrian Carnica honey bee populations, managed by two different breeding associations, the ACA and the ZAC!, using comprehensive pedigree information.

## 2. MATERIALS AND METHODS

### 2.1. Pedigree data and pedigree nomenclature

Pedigree recording in honey bees slightly differs from other livestock species due to their specific nature of reproduction and coancestry within colonies. As a consequence, a pedigree nomenclature, that illustrates both the queen pedigree and the colony pedigree (Figure 1), has been developed (Ruttner 1972). The queen, referred to as 1a, descends from her mother (2a) that was mated to the drones of daughters (2b) from a so-called sire colony with a 6a queen. On colony level, the workers of the 1a colony descent from their mother, the 1a queen, and paternally from drones of different full sib 1b daughters out of a 4a queen. In general, the 4a colony, i.e., the 4a queen and its mating partners, also are termed sire colony of 1a. In the database systems of BeeBreed and BeeData, pedigree information is stored by the input of the registry number of the 1a queen, the registry number of the mother 2a, and the registry number of the 4a queen. In order to generate a queen pedigree on population level, which is the basis for the analysis of gene pool and genetic diversity in honey bee breeding populations, it was necessary to generate the so-called 1b dummy sires (drone-producing full sib daughters of the 4a queen) and further on to link them to their respective 2a mating partners throughout the population pedigree.

The pedigree files were built stepwise, finally leading to dams and sires without known parents, so-called founder dams and founder sires. In total, 43.854 individual records from the years of birth of 1953 to 2020 were available for the ACA population and 17.526 records from the years of birth of 1997 to 2020 for the ZAC! population. In a first step, we generated 1b dummy individuals out of the 4a mothers and linked them to the 1a and 2a queens. The resulting queen pedigree,

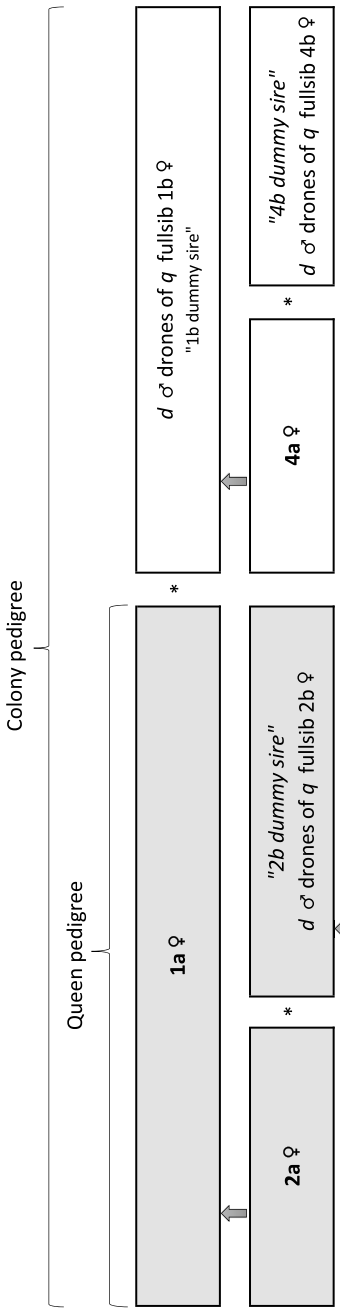


**Figure 1.** Nomenclature of a honey bee colony pedigree (gray boxes indicate the queen pedigree; * = mating; arrow = generation step; ♀ = queen symbol; ♂ = drones symbol; *q* = number of full sib drone producing queens on mating station; *d* = average number of mated drones per queen; bold letters indicate IDs recorded in data base systems of breeding organizations).

defined by the ID triplet 1a (individual queen), 2b (dummy sire), and 2a (dam), was further complemented by the founder animals, i.e., individuals with missing sires or dams, respectively. At this stage, the ACA pedigree comprised 47.196 individual records and the ZAC! pedigree 18.883 individual records.

The reference populations, i.e., the actual breeding populations, were defined by the following assumptions: queens born in 2019 and 2020 and sires (1b queens) of mothers (4a) born in 2016 or later. These restrictions resulted in an active breeding population of 2.675 animals (30 sires (1b); 2.645 females (1a)) in the ACA population and 1.286 queens (26 sires (1b); 1.260 females (1a)) in the ZAC! breeding population. Pedigrees for both reference populations were built up generation by generation using the raw pedigree data. The finally extracted ACA pedigree comprised 4.079 individual records and reached a maximum pedigree length of 23 generations with the oldest entry in the year of 1985. The pedigree of the ZAC! population consisted of 1.830 individual records with its oldest registration in 1995. As a consequence of this method, pedigree information is available for each queen and generation.

### 2.2. Methods

Pedigree data were used to calculate the following measures of genetic variability: effective numbers of founders ($f_e$), ancestors ($f_a$), and founder genome equivalent ($n_g$), as described by Boichard et al. (1997). The effective number of founders (Eq. 1) is defined by the number of equally contributing founders that would be expected to generate the observed genetic diversity in the reference population (Lacy 1989).

$$f_e = \frac{1}{\sum_{k-1}^{f} q_k^2} \tag{1}$$

Here, $f$ represents the number of founders, and $q_k$ is the genetic contribution of the $k$th founder to the reference population. In the case of equal founder contributions, the effective number of founders will be the total number of founders. When breeding animals are used in an unbalanced way, the effective number of founders can be smaller than the total number of founders. Additionally, the effective number of ancestors ($f_a$) takes the occurrence of genetic bottlenecks in pedigrees into account (Boichard et al. 1997). It represents the minimum number of equally contributing ancestors, i.e., influential breeding animals (they may be founders or not) that are necessary to explain the genetic variation in the reference population. The expected marginal contribution ($p_j$) of each ancestor ($_j$) represents its expected genetic contribution without contributions from other (related) ancestors. The ancestors get selected iteratively, according to their marginal contribution ($p_k$). $p_k$ is defined as the $k$th ancestor contribution not yet accounted for by the ($k$-1) ancestors already chosen. Then, this procedure will be iterated, until all contributing ancestors are defined (Eq. 2).

$$f_a = \frac{1}{\sum_{k-1}^{f} p_k^2} \tag{2}$$

$p_k$ is the marginal genetic contribution of the $k$th ancestor not accounted for by the previous $(k-1)$ ancestors, whereas $f$ is the number of ancestors.

The founder genome equivalent ($n_g$) accounts for unequal founder contributions, pedigree bottlenecks, and random loss of alleles due to genetic drift (Lacy 1995) (Eq. 3). This measure of genetic variability is estimated by so-called gene-dropping procedures, where two alleles per founder get segregated iteratively through the pedigree in order to determine the allele retention.

$$n_g = \frac{1}{\sum_{k=1}^{f} \left( \frac{p_k^2}{r_k} \right)} \tag{3}$$

$f$ is the number of founders, $p_k$ is the genetic contribution of founder $k$, and $r_k$ is the allele retention.

Whereas the summing-up of founder contributions ($q_k$) and ancestor contributions ($p_k$) can correctly be performed on basis of a queen pedigree, the estimation of relationship and further

on of inbreeding will be affected by honey bee specific nature of reproduction. On the maternal side, 50% of genes will be transmitted to the next generation resulting in a relationship coefficient ($R$) of 0.5 between dam and daughter. On the paternal side, the relationship between offspring deviates due to polyandry. Virgin queens on a mating station will have the opportunity to mate different drones, which originate from $q$ full sib drone producing queens. As a result, relationship and further on inbreeding will be overestimated when using a honey bee pedigree including single dummy sires on the paternal side of pedigree triplets (comp. Supplementary file 1). Bienefeld et al. (1989) were able to show that honey bee relationship coefficients between offspring queens can be approximated by a function of average number of drones $d$ mating a dam and the number $q$ of drone-producing queens on a mating station. Instead of an expected relationship between full sibs of $R=0.5$ in diploid organisms, the approximation in honey bees results in a coefficient of 0.401 assuming $q=12$ and $d=8$ (Bienefeld et al. 2007; Supplementary file 1). In this paper, we calculated inbreeding according to the queen pedigree build up upon "2b dummy sires," which results in overestimated estimates, that further can be used only to compare differences between the studied queen populations of ACA and ZAC!.

Inbreeding coefficients were calculated for individuals considering five generations ($F_5$) according to the method of Meuwissen and Luo (1992). The quality of the underlying pedigree information is described by the complete generation equivalent and the completeness of the pedigree. The complete generation equivalent is computed as the sum over all known ancestors of the terms computed as the sum of $(1/2)^n$ where $n$ is the number of generations separating the individual to each known ancestor (Maignel et al. 1996). We further described the gene pool of both breeding populations by means of individual founder contributions and ancestor contributions, which can be derived from the average relatedness matrix (AR) by the relationship of $\text{founder}_k$/$\text{ancestor}_k$ to $\text{individual}_i$ of the reference population (Gutiérrez and Goyache 2005). The genealogical structure, i.e., the structure of maternal lines within populations, was studied and graphically visualized using the software Pedigraph (Garbe and Da 2003). The genetic diversity parameters $f_e$, $f_a$, inbreeding and pedigree completeness, and individual founder and ancestor contributions were calculated using the software package ENDOG (Gutiérrez and Goyache 2005). The effective number of founder genomes $n_g$ was determined with the PMX software package (Lacy et al. 2012). Statistical analyses and the processing of raw pedigree data were performed using the software package SAS studio 3.8 (SAS 2018).

## 3. RESULTS

The ACA reference population comprised in total 2.675 queens, from which 2.261 queens descended from 110 sires and 429 dams. For 414 queens, the sire was not known. Of these 414 queens, 29 animals also had an unknown dam. Thus, 414 queens within the reference/testing population of 2019 and 2020 represented founder animals, which can be regarded as foreign Carnica gene-flow, i.e., integration of foreign Carnica genetics into the ACA studbook. A similar situation was observed within the ZAC! reference population. Here, 1.478 queens originated from 29 sires and 162 dams. For 352 queens, the sire was not known. They might originate from other Carnica breeders, other Carnica breeding associations, or unbred, wild Carnica colonies.

The quality of the pedigree information can be described by the maximum number of generations, the completeness of pedigree information per generation (Figure 2), and by the so-called generation equivalent, which is the average number of fully documented generations (Table I). The pedigree completeness, as illustrated in Figure 2, showed differences between the two breeding associations from generation one to generation five. In the second generation, 80.6% of ancestors were known in the ACA population. For the ZAC! population, this percentage comprised 63.4%. The decline in pedigree documentation was faster for the ZAC! studbook than for the ACA studbook, but in generation five, the percentage of known ancestors (ca. 35%)

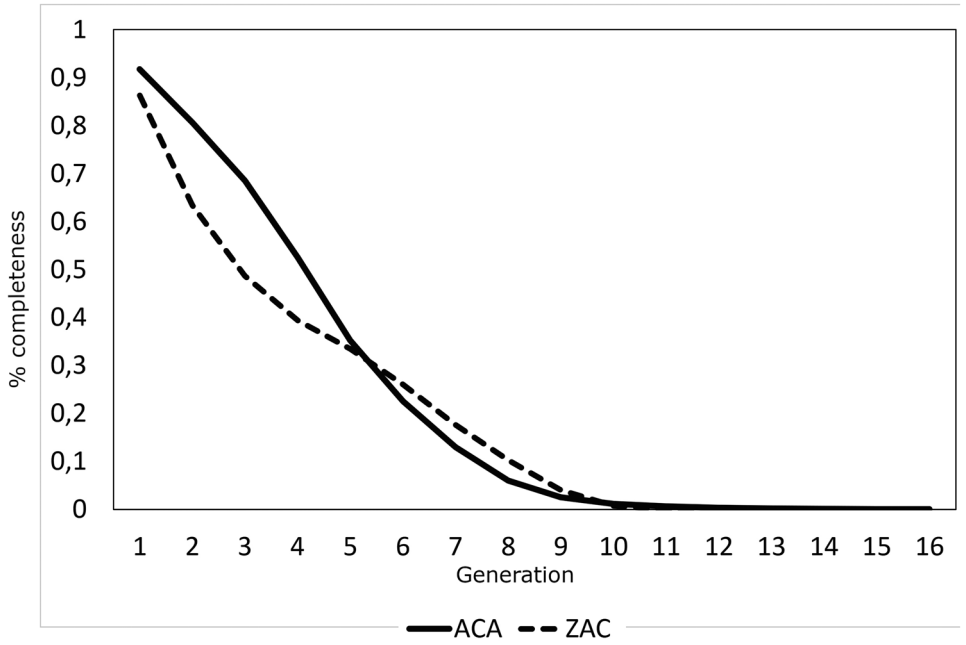


**Figure 2.** Percentage of known ancestors per generation for the ACA and ZAC! Carnica honey bee pedigrees.

was almost the same for both breeding associations. This difference in pedigree documentation has been confirmed by the mean generation equivalent, where active ACA queens provide on average 3.8 generations of fully documented pedigrees and active ZAC! queens provide on average only 3.3 generations of fully documented pedigrees.

For the reference populations of the ACA and the ZAC!, the mean inbreeding coefficient calculated on the basis of five generations reached 3.0% or 2.1%, respectively (Table I).

Considering only inbred animals of both reference populations (53.6% of the ACA and 40.2% of the ZAC! population were inbred), the mean generation equivalent of both samples increased to five fully documented generations on average and the mean inbreeding level was about two times higher than in the original reference populations (Table II).

The indice effective number of founders ($f_e$), effective number of ancestors ($f_a$), and effective number of founder genomes ($n_g$) are used as indicators for the genetic diversity. In total, the ACA reference population is based on 1.015 founder animals. The reference population also included 414 founders out of in total of 1.015 founder animals. These 414 founder queens thus represent recent immigration of foreign Carnica genetics into the ACA population of breeding queens. This novel immigration gene pool can be quantified by 9.2%. The $f_e$ value for the ACA reference population

**Table I** Mean inbreeding values ($F$ = inbreeding based on 5 generations and complete generation equivalent for the two Carnica honey bee populations of the ACA and the ZAC!

| Population | Reference population | Pedigree size | Max. number of generations | Generation equivalent | $F_5$ |
|---|---|---|---|---|---|
| ACA | 2675 | 4079 | 23 | 3.75 | 2.99 ± 5.03 |
| ZAC! | 1286 | 1830 | 13 | 3.29 | 2.13 ± 3.97 |

comprised the number of 71, and 32 founders were necessary to explain 50% of the ACA gene pool (Table III). The $f_a$ value was 30, and 13 influential breeding animals accounted for 50% of the gene pool. The founder genome equivalent ($n_g$) comprised a value of 18.68. In Figures 3 and 4, the contribution of single founders and ancestors to the ACA gene pool is illustrated (Supplementary file 2). In addition, it was possible to assign founder animals to the country of their origin for the ACA population. 231 founder animals of in total 1.015 founders were re-imported from the German breeding associations DIB (Deutscher Imkerbund) or AGT (Arbeitsgemeinschaft Toleranzzucht), respectively. Currently, these foundation breeding queens account for 53.5% of the current ACA gene pool.

For the ZAC!, the novel Carnica immigration gene pool comprised 13.8% originating from 352 "neo" founder individuals among the reference population. The diversity indices of the ZAC! population were generally higher with $f_e = 125$, $f_a = 48$, and $n_g = 21.60$. Here, 50 founders or 21 influential breeding animals, respectively, contributed to 50% of the ZAC! Carnica gene pool (Table III, Figures 3 and 4, and Supplementary file 3).

The genealogical structure, as defined by maternal lines, was extracted for both breeding populations. The ACA population is characterized by in total 291 maternal lines, of which 133 lines were represented by 2.645 active breeding animals. The other 158 lines died out, i.e., did not contribute living animals to the current breeding population (Supplementary file 4). Of the 133 breeding lines, 74 genealogical strains (Table V) can be regarded as a result of novel genetic immigration (comp. foreign Carnica gene-flow in Supplementary file 2, Table V).

The remaining 59 maternal lines thus represent the basic genealogical structure of the ACA breeding population (Table IV). The majority of queens ($n = 2.319$) were members of these basic maternal lines, 22 of these lines could be assigned to German founder queens (1.338 living animals), one line had a Croatian background (39 living animals), and 36 maternal lines were of Austrian origin (942 living animals) (Table IV).

The integration of foreign Carnica genetics into the ACA population of breeding queens was demonstrated by 74 novel immigration lines, which again could be split in a group of German origin (48 lines with 258 active queens) and an Austrian group containing 26 introgression lines with 68 living queens (Table V). The family sizes (number of living queens per line) were highest in the German genealogies (mean family size = 60.8 queens per line) and lowest in the Austrian maternal lines (mean family size = 26.2 queens per line) (Table IV). Also, a similar structure was obtained for the novel introgression lines (Table V).

The ZAC! breeding population was genealogically structured by 142 maternal lines in total, for 102 lines living offspring could be determined (Supplementary file 5). Here, a novel immigration pool was observed too, consisting of 56 breeding lines that were recently added and that comprise a total of 450 living queens. The long-term breeding stock harbored 46 different of the so-called traditional lines represented by a total of 836 family members. The mean family sizes were smaller than overserved in the ACA population, as they comprised on average 8 living queens per line within the introgression lines and 18.2 living queens per line in the original maternal lines (Table VI).

**Table II** Mean inbreeding values ($F_5$ and complete generation equivalent (average number of complete generations)) for the two inbred Carnica honey bee populations of the ACA and the ZAC!

| Population | Reference population | Inbred reference population | Generation equivalent | $F_5$ |
|---|---|---|---|---|
| ACA | 2675 | 1435 | 4.86 | 5.59 ± 5.72 |
| ZAC! | 1286 | 517 | 5.48 | 5.32 ± 4.73 |

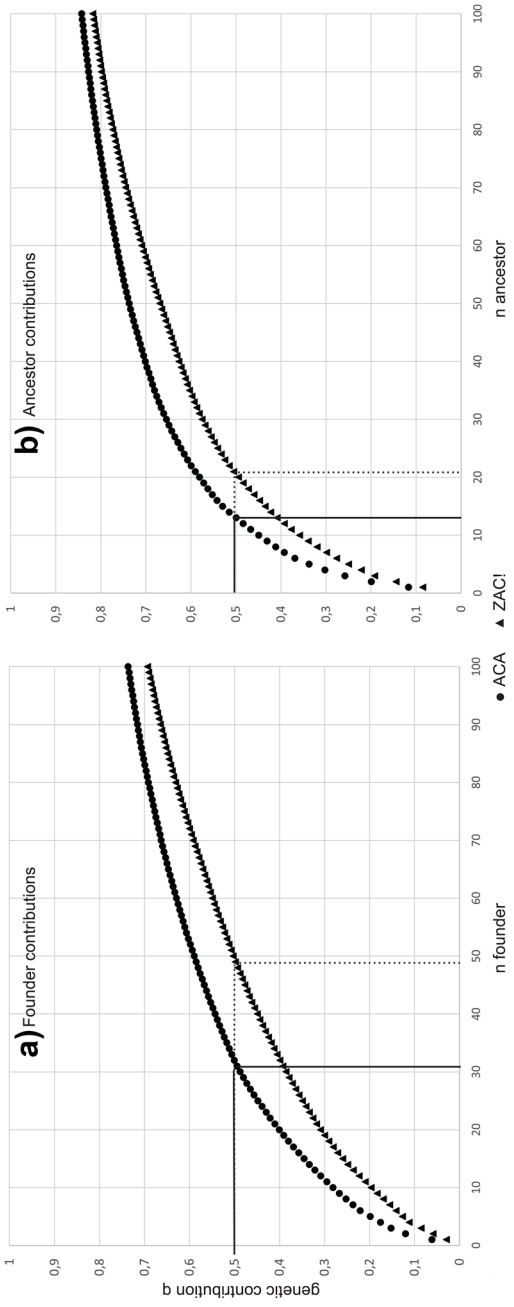


**Figure 3.** Genetic contributions of the first 100 founder/ancestors to the gene pool of the ACA and ZAC! Carnica breeding populations.

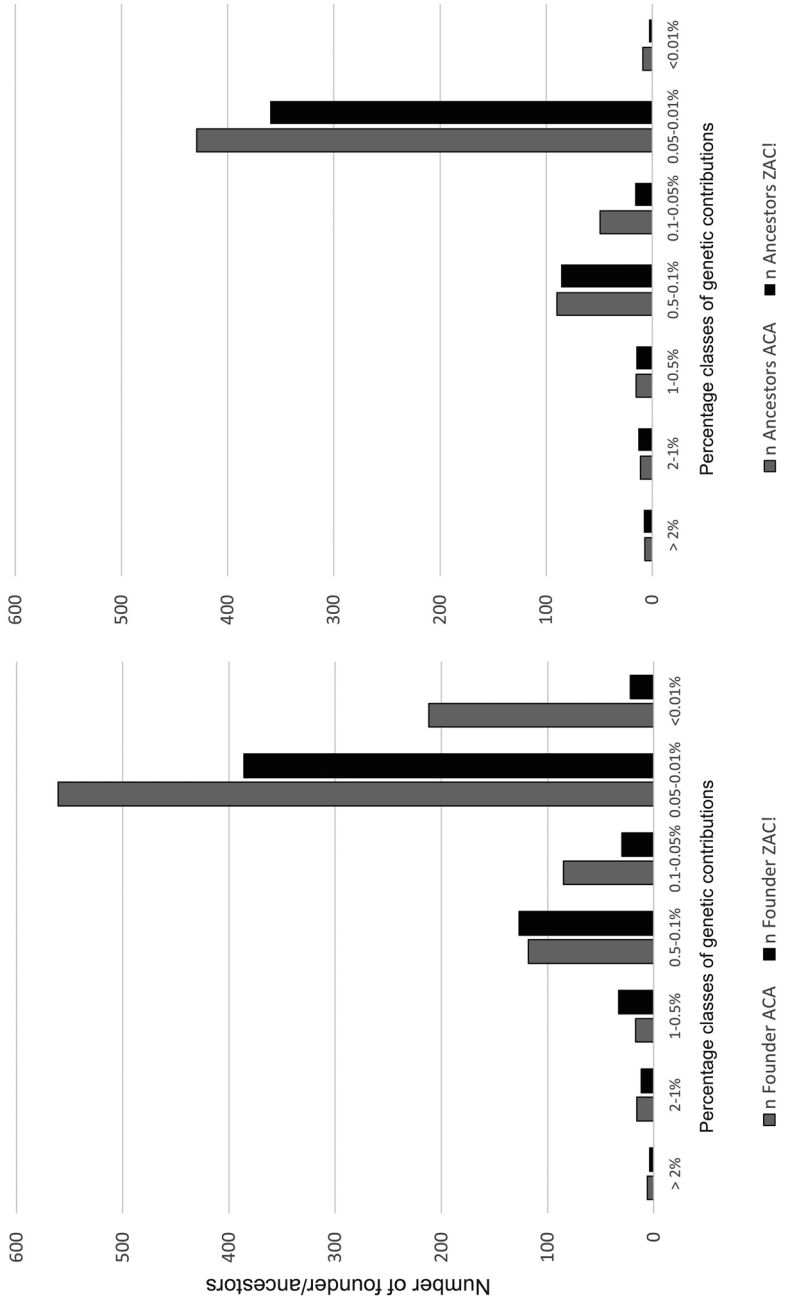


**Figure 4.** Number of founder/ancestors within classes of genetic contribution: more than 2%; 1–2%, contributing 0.5–1%, 0.1–0.5%, 0.05–0.1%, and 0.01–0.05%, less than 0.01%; to the gene pool of the ACA and ZAC! Carnica breeding population.

**Table III** Diversity indices effective number of founders ($f_e$), effective number of ancestors ($f_a$), and effective number of founder genomes ($n_g$), number of founders, number of founders contributing to 50% or 75% of gene pool, respectively (founders$_{.50/0.75}$), number of ancestors, number of ancestors contributing to 50% or 75% of gene pool, respectively (Ancestors$_{.50/0.75}$), for the two Austrian Carnica honey bee populations

| Population | Founders | Founders$_{.50/0.75}$ | Ancestors | Ancestors$_{.50/0.75}$ | $f_e$ | $f_a$ | $n_g$ |
|---|---|---|---|---|---|---|---|
| ACA | 1015 | 32/107 | 610 | 13/55 | 71 | 30 | 18.68 |
| ZAC! | 624 | 50/125 | 501 | 21/72 | 125 | 48 | 21.60 |

## 4. DISCUSSION

*A. m. carnica* can be regarded as a transboundary breed with multiple subpopulations spread across the European and North American continents. In Europe, the international breeding program "Beebreed" exists, which is administered by the LIB and currently supports Carnica breeding associations of 18 European countries. As usual in honey bee breeding, the Carnica breeding programs are characterized by semi-open studbooks, i.e., they allow the use of Carnica breeding animals, which did not belong to the own breeding population. In this study, we were able to quantify such introgression of foreign Carnica genetics, which comprised for the ACA studbook 53.5% of the entire gene pool represented by 71 maternal lines of a total of 133 lines. For the ZAC! population, the origin of the founders could not be traced since they stem from a split-off the ACA breeding population in 2002, but within the active breeding populations of both Austrian breeding associations, the so-called "neo"-founder animals were detected. The genetic contribution of those founders (foreign origin or no pedigree available) to the ACA gene pool was 9.2% (414 founder queens) and 13.8% (352 founder queens) for the ZAC! gene pool.

Such constant genetic admixture nominally decreases the overall coancestry within a population resulting in underestimated inbreeding levels, as the inbreeding coefficient is directly related to the amount of pedigree documentation available. For the ACA reference population, a mean $F_5$ coefficient of 3.0% and for the ZAC! reference population a $F_5$ of 2.1% were estimated (overestimated values, as underlying paternal pedigree relationship was not approximated using the method of Bienefeld et al. 2007). In order to determine the lower border of genetic variability within the two Austrian Carnica breeding populations, we limited the reference populations to animals with a full pedigree record (the mean generation equivalent is approximately 5). This resulted in 1.435 active breeding queens (53.6%) of the ACA breeding population and 517 active breeding queens (40.2%) of the ZAC! breeding population. As expected, we observed an increase of inbreeding for about two times ($F_5 = 5.6\%$ in the ACA and $F = 5.32\%$ in the ZAC! breeding population). In general, the estimation of inbreeding in honey bees can be biased by several factors: incorrect pedigree documentation, deviation of paternal ancestry due to reduced levels of isolation of

**Table IV** Descriptive statistics for active queens of 59 traditional maternal lines (long-term families) which structure the ACA Carnica honey bee population (*n*, number; *Min. and Max.*, min. and max. number of living queens per line)

| Maternal lines | *n* lines | *n* queens | Mean *n* queens per line | Min. | Max. |
|---|---|---|---|---|---|
| Maternal lines of German origin | 22 | 1338 | 60.82 ± 77.28 | 1 | 290 |
| Maternal lines of Austrian origin | 36 | 942 | 26.17 ± 37.57 | 1 | 208 |
| Maternal line of Croatian origin | 1 | 39 | - | - | - |
| Sum of maternal lines | 59 | 2319 | 39.31 ± 57.42 | 1 | 290 |

**Table V** Descriptive statistics for active queens of newly introduced foreign lines (hereby named introgression lines) that structure the ACA Carnica honey bee population and for all 133 queen lines within the ACA population (*n*, number; *Min. and Max.*, min. and max. number of living queens per line)

| Introgression lines | *n* lines | *n* queens | Mean *n* queens per line | Min. | Max. |
|---|---|---|---|---|---|
| Introgression lines of German origin | 48 | 258 | 5.38 ± 5.61 | 1 | 28 |
| Introgression lines of Austrian origin | 26 | 68 | 2.62 ± 2.58 | 1 | 11 |
| Sum of introgression lines | 74 | 326 | 4.41 ± 4.93 | 1 | 28 |
| Sum of all lines within ACA | 133 | 2645 | 19.89 ± 42.01 | 1 | 290 |

controlled mating stations, and average approximation methods of paternal relationship according to Bienefeld et al. (2007) or Brascamp and Bijma (2014). Pedigree errors are a common fact in livestock breeding and are mostly due to errors in documentation or identification of individuals. In dairy cattle breeding programs, pedigree errors within a range of 4–13% have been reported by Visscher et al. (2002) and Leroy et al. (2012); in goat, sheep, and dog breeds, error rates can reach up to 22%, followed by pig breeds with maximum error rate up to 37% (Zhang et al. 2020). In honey bee breeding, artificial insemination and island mating station (f.e. in Germany) will be able to provide high levels of accuracy in paternal ancestry. In Austria, mating stations are located generally in isolated alpine regions and are therefore surrounded by natural borders. The mating success rate is estimated > 75% by the Austrian breeding associations. In general, pedigree-based inbreeding ($F_{PED}$), i.e., the probability of alleles that are identical by descent (IBD), will deviate from comparable genomic based information (Kardos et al. 2015). Since the availability of large scale genomic data, it became feasible to derive reliable genomic IBD estimates like runs of homozygosity (ROH) and subsequently the $F_{ROH}$ statistics (McQuillan et al. 2008, Purfield et al. 2017; Ceballos et al. 2018). In a recent study, Gmel et al. (2023) presented ROH statistics for Western honey bees. The authors were able to show that in honey bees, the correlation between $F_{PED}$ and $F_{ROH}$ comprised -0.22, indicating a low to non-existing relationship between both measures of inbreeding. In livestock breeds, correlation coefficients between the two parameters also range at a lower to medium scale ($rF_{ROH\_}F_{PED}$ from 0.18 to 0.70; Purfield et al. 2017). Although pedigree-based estimates of inbreeding do not measure IBD at the same precision level as genomic marker-based methods, the rather conservative approach of pedigree analysis via founder and ancestor contributions still can provide valuable information on the performance and management of breeding programs and the composition of the respective gene pools.

The genetic diversity indices $f_e$, $f_a$, and $n_g$ are capable to determine unbalanced breeding practices, occurrence of genetic (pedigree) bottlenecks, and genetic drift in the respective population histories. The values obtained in this study (ACA/ZAC!: $f_e = 71/125$; $f_a = 30/48$;

**Table VI** Descriptive statistics for active queens from traditional (maternal) and newly introduced foreign lines (introgression lines) that structure the ZAC! Carnica honey bee population (*n*, number; *Min. and Max.*, min. and max. number of living queens per line)

| Queen lines | *n* lines | *n* queens | Mean *n* queens per line | Min. | Max. |
|---|---|---|---|---|---|
| Maternal lines | 46 | 836 | 18.17 ± 11.47 | 1 | 44 |
| Introgression lines | 56 | 450 | 8.04 ± 5.78 | 1 | 40 |
| Sum of all lines within ZAC! | 102 | 1286 | 12.61 ± 10.12 | 1 | 44 |

$n_g = 18.7/21.6$) demonstrated a tendency of genetic loss due to an excessive use of single breeding animals within the ACA breeding program and a decrease of genetic diversity within the ZAC! population that mostly is due to smaller population size. In comparison, horse breeds with limited population size ($n$ from 565 to 2.808) and closed studbooks for more than 20 generations, such as Lipizzans (Zechner et al. 2002), Haflingers (Druml et al. 2016), and working horses (Druml et al. 2009; Janssens et al. 2010), exhibit $f_e$ values between 31 and 117 and $f_a$ values between 12 and 29. In canine populations, which are characterized by high phenotypical, breeding census, and founder gene pool variation, Wijnrocx et al. (2016) presented an overview on pedigree parameters for 23 Belgian dog breeds. Here, $f_e$ varied from 302 to 3 (mean = 126.8), $f_a$ varied from 157 to 3 (mean = 57.9), and $n_g$ ranged from 78.1 to 1.2 (mean = 27.6). The diversity indices measured for the two Austrian honey bee breeding populations were considerably lower and underline the limitations within the Austrian Carnica gene pools of breeding queens.

The commonly applied strategy of supporting admixture by genetic immigration can have positive, short-term effects on the genetic diversity of the studied Carnica honey bee breeding populations. In long term, the additional imported gene pool (9.2% of the ACA and 13.8% of the ZAC! gene pool) will constantly decrease the genetic contribution of founder queens present in small proportions and eventually eliminate them. Together with unequal founder and ancestor contributions, respectively, in an EBV-assisted mass selection program, this tendency can get enforced. Taking conservation genetics into account, minimal founder-gene proportions are essential for the diversity within a gene pool. In the ACA breeding population, 589 founder animals constitute 50.8% of the gene pool with a mean contribution of 0.09% per animal. In the ZAC! population, 242 founders contribute to 41% of segregating alleles with a mean contribution of 0.17% per individual. As a consequence, the strengthening of founder gene proportions at low frequencies could assist in maintaining the genetic diversity in both honey bee breeding populations and prevent the subsequent, constant loss of rare alleles in the future. A simple approach on how to minimize coancestry and promote rare genetic diversity is the balanced build-up of maternal genealogies. Currently, these structures are highly unbalanced, as in the ACA 25 lines out of 59 consist of less than 15 family members (mean = 6.8). The remaining 34 lines comprise on average 63.2 family members. A comparable situation can be observed in the ZAC! population (19 lines with on average 7.8 members; 27 lines with on average 25.4 members). Thus, a strategic reinforcement of the small families would strengthen the rare alleles within the Carnica honey bee gene pool of both breeding populations and contribute to their conservation.

## 5. CONCLUSION

In this pedigree study on two Austrian honey bee breeding populations, we were able to describe the gene pool and the genetic diversity by means of founder and ancestor contributions, genetic diversity indices, and genealogical structures. Due to limited size of breeding population within the ZAC! breeding association, and unbalanced use of breeding animals within the ACA breeding association, both populations exhibit a moderate loss of genetic diversity comparable to mammal livestock with small or limited population size.

However, the presence of two differentiated breeding populations of *A. m. carnica* in Austria originating of a same founder stock and the presence of sufficient genealogical structure represent valuable genetic resources that might be of use in future breeding plans supporting the maintenance of gene pool variability. Additional pedigree analyses of honey bee populations of other breeding programs or subspecies would help to put diversity indices and founder gene pool architecture into a species-specific context.

## SUPPLEMENTARY INFORMATION

The online version contains supplementary material available at https://doi.org/10.1007/s13592-023-00999-w.

## AUTHOR CONTRIBUTION

All authors contributed to the study conception. Material preparation and data collection were performed by Michael Rubinigg, Christian Boigenzahn, Anselm Putz, Martin Kärcher, and Karl Neubauer. The analyses were performed by Thomas Druml. The first draft of the manuscript was written by Thomas Druml, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

## FUNDING

This study was funded by the Austrian "Bundesministerium für Land- und Forstwirtschaft, Regionen und Wasserwirtschaft (BML)" and the "Biene Österreich" (Grant number: 101720).

## DATA AVAILABILITY

The datasets of this study are available from the corresponding author on a science-based request.

## CODE AVAILABILITY

Not applicable.

## DECLARATIONS

**Ethics approval** Not applicable.

**Consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Competing interests** The authors declare no competing interests.

## REFERENCES

Aberle K, Wrede J, Distl O (2003) Analysis of the population structure of the Black Forest draught horse. Berl Munch Tierarztl Wochenschr 116:333–339

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

Boichard D, Maignel L, Verrier E (1997) The value of using probabilities of gene origin to measure genetic variability in a population. Genet Sel Evol 29:5–23

Boigenzahn C (2020) Der österreichische Imkereisektor Wien

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

Brascamp EW, Willam A, Boigenzahn C, Bijma P, Veerkamp, R.F. (2016) Heritabilities and genetic correlations for honey yield, gentleness, calmness and swarming behaviour in Austrian honey bees. Apidologie. https://doi.org/10.1007/s13592-016-0427-9

Ceballos FC, Joshi PK, Clark DW, Ramsay M, Wilson JF (2018) Runs of homozygosity: windows into population history and trait architecture. Nat Rev Genet 19:220–234

Cunningham EP (1991) The genetics of thoroughbred horses. Sci Am 264:91–98

Dickson F, Lush JL (1933) Inbreeding and the genetic history of the Rambouillet sheep. Am J Hered 24:19–33

Druml T, Baumung R, Sölkner J (2009) Pedigree analysis in the Austrian Noriker draught horse: genetic diversity and the impact of breeding for coat colour on population structure. J Anim Breed Genet 126:348–356

Druml T, Sauer K, Elsbacher J, Grilz-Seger G, Brem G (2016) Analysis of the gene pool, the genetic diversity and inbreeding ratios of Austrian Haflinger Population. Züchtungskunde 88(5):379–394

Gandini GC, Bagnato A, Miglior F, Pagnacco G (1992) Inbreeding in the Italian Haflinger horse. J Anim Breed Genet 109:433–445

Garbe J, Da Y (2003) Pedigraph, a software tool for the graphical visualization of large complex pedigrees. Final abstracts guide, p 293. Plant and Animal Genome XI, San Diego, CA, January 11–15

Gmel AI, Guichard M, Dainat B, Williams GR, Eynard S, Vignal A, Servin B, Beestrong Consortium, Neuditschko M (2023) Identification of runs of homozygosity in Western honey bees (Apis mellifera) using whole-genome sequencing data. Ecol Evol 13:e9723

Gutiérrez JP, Goyache F (2005) A note on ENDOG: a computer program for analysing pedigree information. J Anim Breed Genet 122:172–176

Hoppe A, Du M, Bernstein R, Tiesler FK, Kärcher M, Bienefeld K (2020) Substantial genetic progress in the international Apis mellifera carnica population

since the implementation of genetic evaluation. Insects 11:768

Janssens S, Stinckens A, Schroyen M, Peeters L, De Keyser K, De Wael R, Lamberigts C, Luyten T, Ons E, Buys N (2010) Genetic diversity in the Belgian Draught Horse breed as revealed by pedigree analysis and molecular marker data. Anim Genet 41(Suppl. 2):205–206

Kardos M, Luikart G, Allendorf FW (2015) Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. Heredity 115:63–72

Lacy RC (1989) Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. Zoo Biol 8:111–123

Lacy RC (1995) Clarification of genetic terms and their use in the management of captive populations. Zoo Biol 14:565–577

Lacy RC, Ballou JD, Pollak JP (2012) PMx: software package for demographic and genetic analysis and management of pedigreed populations. Meth Ecol Evol 3:433–437

Leroy G, Danchin-Burge C, Palhiere I, Baumung R, Fritz S, Mériaux JC, Gautier M (2012) An ABC estimate of pedigree error rate: application in dog, sheep and cattle breeds. Anim Genet 43:309–314

Lodesani M, Costa C (2003) Bee breeding and genetics in Europe. Bee World 84(2):69–85

Lush JL (1947) Family merit and individual merit as bases of selection. Am Naturalist 81(241–261):362–379

Maignel L, Boichard D, Verrier E (1996) Genetic variability of French dairy breeds estimated from pedigree information. Interbull Bull 14:49–54

McCluer JW, Vandeberg JL, Read B, Ryder OA (1986) Pedigree analysis by computer simulation. Zoo Biol 5:147–160

McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS et al (2008) Runs of homozygosity in European populations. Am J Hum Genet 83:359–372

Meuwissen TI, Luo Z (1992) Computing inbreeding coefficients in large populations. Genet Sel Evol 24:305–313

Moreaux S, Verrier E, Ricard A, Meriaux JC (1996) Genetic variability within French race and riding horse breeds from genealogical data and blood marker polyphormisms. Genet Sel Evol 28:83–102

Purfield DC, McParland S, Wall E, Berry DP (2017) The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. PLoS ONE 12:e0176780

Ruttner F (1956) The mating of the honey bee. Bee World 37:3–15

Ruttner H (1972) Technische Empfehlung zur Methodik der Leistungsprüfung von Bienenvölkern. Internationales Symposium, Paarungskontrolle und Selektion bei der Honigbiene, Lunz am See, Österreich, Apimodia Verlag, Bukarest 103–107

Ruttner F (1988) Biogeography and taxonomy of honey bees. Springer Verlag, Berlin

Ruttner F (1996) Zuchttechnik und Zuchtauslese bei der Biene. Ehrenwirt Verlag GmbH, München

SAS (2018) SAS/IML 3.8 User's Guide. SAS Institute Inc., Cary, NC

Visscher PM, Woolliams JA, Smith D, Williams JL (2002) Estimation of pedigree errors in the UK dairy population using microsatellite markers and the impact on selection. J Dairy Sci 85:2368–2375

Wallberg A, Han F, Wellhagen G, Dahle B, Kawata M, Haddad N, Simões ZLP, Allsopp MH, Kandemir I, De la Rúa P (2014) A worldwide survey of genome sequence variation provides insight into the evolutionary history of the honey bee Apis mellifera. Nat Genet 46(10):1081–1088

Wijnrocx K, Francois L, Stinckens A, Janssens S, Buys N (2016) Half of 23 Belgian dog breeds has a compromised genetic diversity, as revealed by genealogical and molecular data analysis. J Anim Breed Genet 133:375–383

Willam A, Essl A (1993) Zuchtwertschätzung für die Honigleistung (Apis mellifera carnica) in Österreich. Die Bodenkultur 44(3):14

Wright S, McPhee HC (1925) An approximate method of calculating coefficients of inbreeding and relationship from livestock pedigrees. J. Agr. Res. 31, 377–383

Zechner P, Sölkner J, Bodo I, Druml T, Baumung R, Achmann R, Marti E, Habe F, Brem G (2002) Analysis of diversity and population structure in the Lipizzan horse breed based on pedigree information. Livest Prod Sci 77:137–146

Zhang J, Song H, Zhang Q, Ding X (2020) Assessment of relationships between pigs based on pedigree and genomic information. Animal 14:697–705. https://doi.org/10.1017/S1751731119002404