

# An abbreviated SNP panel for ancestry assignment of honeybees (*Apis mellifera*)

Nadine C. CHAPMAN<sup>1</sup>, A. Lelania BOURGEOIS<sup>2</sup>, Lorraine D. BEAMAN<sup>2</sup>, Julianne LIM<sup>1</sup>, Brock A. HARPUR<sup>3</sup>, Amro ZAYED<sup>3</sup>, Michael H. ALLSOPP<sup>4</sup>, Thomas E. RINDERER<sup>2</sup>, Benjamin P. OLDROYD<sup>1</sup>

<sup>1</sup>Behaviour and Genetics of Social Insects Lab, School of Life and Environmental Science, University of Sydney, Macleay Building A12, Sydney, NSW 2006, Australia

<sup>2</sup>Honey Bee Breeding, Genetics and Physiology Research Laboratory, USDA-ARS, 1157 Ben Hur Road, Baton Rouge, LA 70820, USA

<sup>3</sup>Department of Biology, York University, 4700 Keele Street, Toronto, Ontario M3J1P3, Canada

<sup>4</sup>ARC-Plant Protection Research Institute, Stellenbosch 7599, South Africa

Received 29 January 2017 – Revised 9 May 2017 – Accepted 23 May 2017

**Abstract** – We examine whether a panel of 37 single nucleotide polymorphisms (SNPs) has the same power as a more expensive panel of 95 SNPs to assign ancestry of honeybees (*Apis mellifera*) to three ancestral lineages. We selected SNPs using allele frequencies, such that poorly performing SNPs were excluded. We find that ancestry assignment is comparable between the two panels. Importation of bee semen from countries where Africanized bees are present into countries where Africanized bees are absent would be facilitated if small proportions of semen derived from Africanized drones can be reliably detected. We used the abbreviated panel to determine if semen from a single Africanized drone could be detected when mixed with the semen of 10, 20 or 40 non-Africanized drones. We found that the use of the 37 SNP test on a mixed sample would fail to detect the contribution of a single Africanized male. It is therefore important that the cadavers of the males contributing semen are individually tested.

**Ancestry informative markers / Single nucleotide polymorphism / Honeybee / *Apis mellifera* / Africanized honeybee**

## 1. INTRODUCTION

Africanized honeybees are hybrids of *Apis mellifera scutellata* from Africa and subspecies from Europe. These bees are widely held to be unsuitable for beekeeping due to their heightened aggression and their strong tendency to abscond (Schneider et al. 2004). Africanized honeybees

are highly invasive, having spread from their introduction point in São Paulo, Brazil, to as far south as Buenos Aires in Argentina (Dietz et al. 1985) and as far north as the southern states of the USA (Jarnevich et al. 2014). Numerous countries have procedures in place to prevent the introduction of African and Africanized bees. Nonetheless, several countries, particularly Australia and New Zealand, could benefit from importing non-Africanized *Varroa*-resistant stock from areas where Africanized bees are extant.

Recently, we developed a test to assign the proportion of ancestry of honeybees (*Apis mellifera*) to three ancestral lineages (African (A), West European (M) and East European (C)) (Chapman et al. 2015) based on 95 single

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s13592-017-0522-6) contains supplementary material, which is available to authorized users.

---

Corresponding author: N. Chapman,  
nadine.chapman@sydney.edu.au  
Manuscript editor: Marina Meixner

nucleotide polymorphisms (SNPs). This test may be used as a biosecurity measure to detect and deny entry of Africanized honeybee queens and potentially semen (Chapman et al. 2015). If a rejection threshold of 15% African ancestry is adopted, then all African and Africanized honeybees (25–100% African ancestry) will be assigned as exceeding the threshold and should be denied importation (Oldroyd and Chapman 2016). The test also enables determination of the proportional ancestry (A, C, M) of any individual honeybee or population of honeybees (Harpur et al. 2015; Chapman et al. 2016). Such testing can be valuable in various contexts. For example, populations derived from a small number of introductions, e.g. on Pacific islands, can be assessed for their degree of admixture. Signals of mixed ancestries in such populations are suggestive of multiple introductions, or secondary introductions. Additionally, SNPs testing may be useful in honeybee conservation programs (Pinto et al. 2014; Ilyasov et al. 2015; Muñoz et al. 2015).

This paper compares the performance of the 95 SNPs panel (Chapman et al. 2015) with that of a reduced panel of 37 ancestry-informative markers (AIMs). The smaller panel provides a significant cost advantage over the 95 SNP panel. Fewer primers are required (approximately 60% cost reduction) and the number of multiplex reactions is reduced (one rather than four; approximately 75% cost reduction). In addition, we determine whether a single Africanized individual can be detected in a vial of semen containing ejaculates of 10–40 non-Africanized males. Due to the small volume of semen obtained from an individual drone (c.a. 1  $\mu$ l), testing of individual ejaculates is not possible. Semen imports may be preferred over queen imports because there is no possibility of importing mites (*Varroa*, *Tropilaelaps*, *Acarapis*) and no need to house bees in a quarantine facility.

## 2. METHODS

### 2.1. Design of the reduced panel

SNPs from the 95 SNP panel were shortlisted for inclusion in the abbreviated panel if they had

an allele frequency of greater than 0.9 in one ancestral lineage (A, C, M) and a frequency of less than 0.1 in the other two ancestral lineages (Chapman et al. 2015), resulting in a shortlist of 46 SNPs. Sequenom ASSAY DESIGN SUITE software was then used to optimize conditions to include as many of the shortlisted SNPs as possible in a single reaction. This resulted in a 41 SNP panel. Four SNPs were subsequently discarded due to poor amplification or inconsistency, leading to a panel of 37 SNPs (see Online Resource S11).

In order to determine whether the abbreviated panel performed as effectively as the 95 SNP panel, we re-typed a subset of 170 individuals previously typed with the 95 SNP panel (Chapman et al. 2015) with the 37 SNP panel using the Agena Sequenom Mass ARRAY MALDI-TOF system at the Australian Genomic Research Facility, with manual review of allele calls. The retyped samples are as follows: 20 *A. m. scutellata* samples, 10 Africanized samples from Brazil, 43 Africanized samples from the USA (TX), 28 *Apis mellifera ligustica* samples, 8 commercial samples from Australia, 5 of which are from the Tasmanian highlands and have high proportions of M ancestry, 29 *Varroa*-resistant samples from the USA, 18 commercial samples from the USA, 11 feral (unmanaged) samples from Australia and 3 *A. m. mellifera* samples (see Online Resource S12). Using the reference samples and methods from Chapman et al. (2015), we assigned the proportion of ancestry of each individual to the three ancestral lineages (A, C, M) using structure (Pritchard et al. 2000).

In order to determine if the two panels result in similar ancestry assignment, we calculated the deviation in lineage assignment for the  $i$ th bee as follows:  $D_i = (C_{i,37} - C_{i,95}) + (A_{i,37} - A_{i,95}) + (A_{i,37} - M_{i,95})$  where, e.g.  $C_{i,37}$  is the proportion of SNPs of the C lineage for the 37 SNP panel for the  $i$ th bee, and  $C_{i,95}$  is the proportion for the 95 SNP panel for the  $i$ th bee.  $D_i$  varies between 0 and 1 for each bee, and  $\sum D_i$  is expected to be 0 if the two panels perform identically. We then tested if this differed from 0 using a  $t$  test.

To determine whether SNP genotypes clustered by population, we performed principal components analysis in SPSS, coding heterozygotes as 0 and homozygotes as either 1 or -1. Missing data was replaced with the mean. Drones were excluded from this analysis ( $N = 4$ ).

We compare the average proportion of successfully typed SNPs with each panel using a related-samples Wilcoxon signed rank test across the 170 bees, to compare the performance of the two panels in terms of completeness of the dataset.

## 2.2. Testing of panel on mixed semen samples

In order to determine whether a semen shipment that contained contributions from Africanized drones in an otherwise non-African sample would be detected and fail our test for Africanization, we typed semen samples that contained varying proportions of semen from known Africanized drones using the 37 SNP panel. Africanized drones were sourced from TX, USA. Non-Africanized ‘commercial’ drones were sourced from the United States Department of Agriculture Honey Bee Breeding, Genetics and Physiology Research Laboratory in Baton Rouge, USA.

We created semen test samples containing ratios of semen contributions of 1:10, 1:20 and 1:40 Africanized to non-Africanized individuals. All semen was taken from each individual and each individual contributed to only one semen test sample. Therefore, semen was collected from 70 males from each non-Africanized colony and from 3 males from each Africanized colony. Four pairs of Africanized/non-Africanized colonies were used as drone sources, resulting in four biological replicates for each ratio. DNA was extracted from 6  $\mu$ l of the mixed semen samples using the Qiagen DNeasy extraction kit, with a final elution step of 60  $\mu$ l.

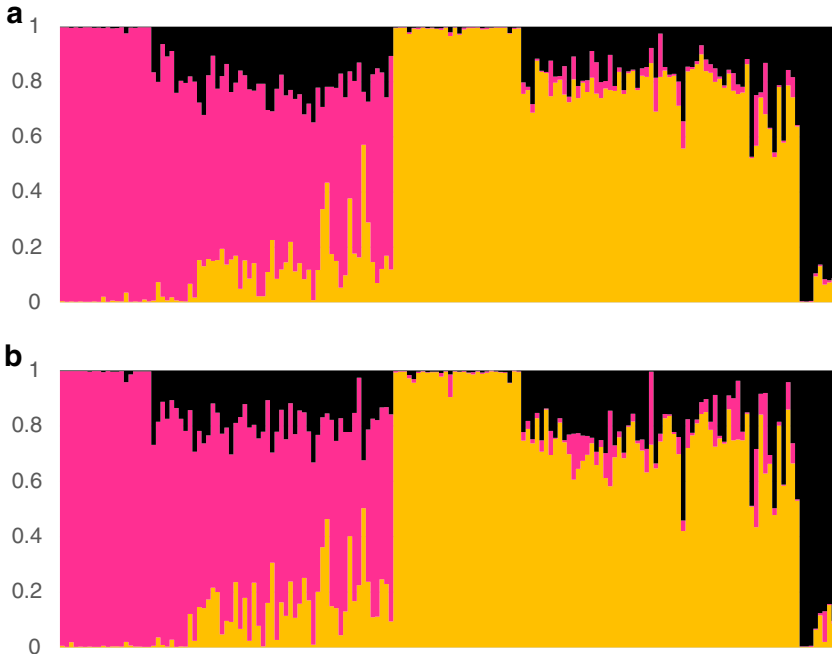
To determine whether all the non-Africanized males collected from each replicate were brothers or non-natal drifters, we extracted DNA from one leg of every non-Africanized drone that contributed semen using Chelex (Walsh et al. 1991) and

genotyped all individuals with five microsatellites (A107, A113, A14, A29, B124 (Estoup et al. 1993) and CSD (Oldroyd et al. 2011)). Based on these genotypes, we selected five natal drones from each replicate, all drifted drones and all Africanized drones, for further analysis. DNA was extracted from two legs from these selected individuals using phenol chloroform in order to produce enough high quality DNA for SNP typing. Only five natal drones from each replicate were SNP genotyped with the 37 SNP panel because drones only inherit DNA from the queen, and five should be sufficient to establish the level of African ancestry in the queen. If a queen has greater than 15% African ancestry, then all semen samples will be rejected on this basis, regardless of whether the contribution of the single Africanized drone included in each sample was detected or not.

Semen samples (12), five natal drones from each replicate (total of 20) and all drifted (25) and Africanized drones (12) were typed with the 37 SNP panel using the Agena Sequenom MassARRAY MALDI-TOF system at the Australian Genomic Research Facility. Using the reference samples and methods from Chapman et al. (2015), we assigned the proportion of ancestry of each individual and each semen sample to the three ancestral lineages (A, C, M) using structure (Pritchard et al. 2000). We determined whether the mixed semen samples and individual drones would be rejected at quarantine. This detection was based on the estimated membership coefficients for each individual or mixed semen sample in each structure cluster (Q). Values of Q (African) above 0.15 were classified as having too much African ancestry and would be rejected at quarantine.

## 3. RESULTS

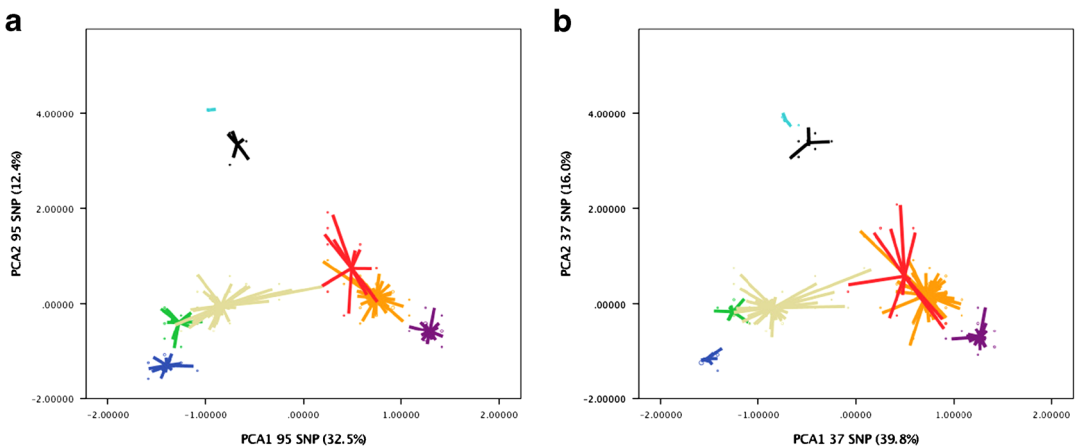
The ancestry assignment of the 37 SNP panel is similar to that given by the 95 SNP panel (see Online Resources SI2; SI3). The mean value of  $D$ , the deviation of proportional ancestry between the two panels, was  $1.8 \pm 0.8\%$ . Although this is significantly different from 0 ( $t_{169} = 2.275$ ,  $P = 0.024$ ; Figure 1), this is unlikely to be of



**Figure 1.** Results of structure analysis for three ancestral lineages (African (A; *pink*), Western European (M; *black*) and Eastern European (C; *yellow*)) using **a** a panel of 95 SNPs (Chapman et al. 2015) and **b** a panel of 37 SNPs for 170 genotyped individuals.

practical importance. The first two principal components explained 32.5 and 12.4% of the variance in the data collected with the 95 SNP panel. The first two principal components explained 39.8 and

16.0% of the variance in the 37 SNP panel (Figure 2). The proportion of SNPs that were typed per individual was significantly higher for the 37 SNP panel ( $99.34 \pm 0.112\%$ ) than the 95



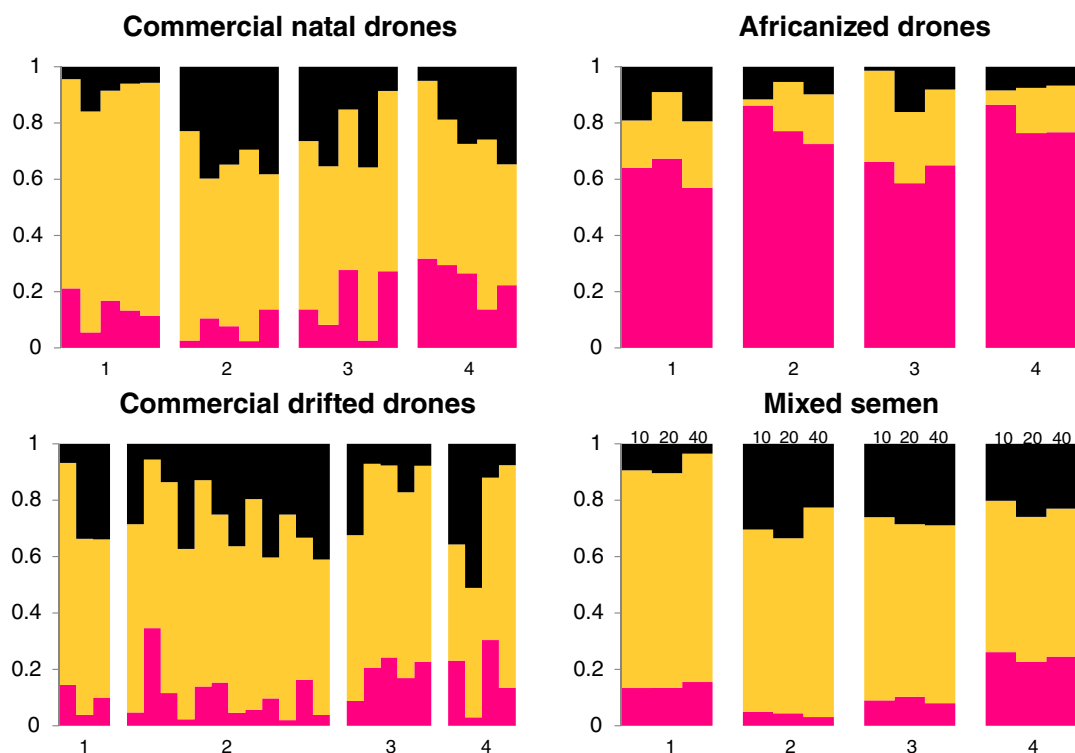
**Figure 2.** Population structure according to the top two principal components for **a** a panel of 95 SNPs and **b** a panel of 37 SNPs for 166 diploid individuals; African (A; *dark blue*), Eastern European (C; *purple*), Western European (M; *light blue*), Africanized bees from Brazil (*green*), Africanized bees from the USA (*khaki*), commercial bees from Australia and the USA (*orange*), unmanaged bees from Australia (*red*) and bees from the Tasmanian highlands, Australia (*black*).

SNP panel ( $88.48 \pm 0.531\%$ ;  $W_{169} = 14,358.500$ ,  $P < 0.001$ ).

All 12 Africanized drones were correctly assigned as being Africanized (African ancestry 56.9–86.5%). Eight of the 20 natal commercial drones (5 natal drones were SNPs genotyped from each of 4 replicates) also exceeded the 15% African threshold. Nine of the 25 drifted drones exceeded the 15% African ancestry threshold. All three mixed semen samples (1:10, 1:20 and 1:40 Africanized:non-Africanized) from replicate 4 exceeded the 15% threshold (African ancestry 22.7–26.1%), as did one sample (1:10 Africanized:non-Africanized) from replicate 1 (African ancestry 15.5%). All other semen samples were below the threshold (African ancestry 3.10–13.4%, Figure 3).

#### 4. DISCUSSION

Ancestry assignment is not differentiable between the 37 SNP panel and the 95 SNP panel. In addition, the 37 SNP panel successfully calls a greater proportion of SNPs than the 95 SNP panel. This improvement, along with a significant cost saving (approximately 60–70%), results in us recommending the use of the 37 SNP panel for honeybee ancestry assignment. Further improvements could be made by including SNPs that are diagnostic of the other ancestral lineages: O (Ruttner 1988; Frank et al. 2000), Y (Frank et al. 2001; Alqarni et al. 2011) and Z (Alburaki et al. 2013). This would perhaps result in lower levels of African ancestry being found in some individuals in commercial populations



**Figure 3.** Proportion of African (A; pink), Western European (M; black) and Eastern European (C; yellow) ancestry in drones collected from commercial US colonies, drones that had drifted from other commercial colonies into the sampled colonies, Africanized drones and mixed semen samples from four replicates. *Numbers above the bars* on the mixed semen sample indicated the number of non-Africanized drones that contributed semen to each mixed semen sample, one Africanized drone contributed to each of the mixed semen samples. Data from each replicate is identified by a *number under the graph*.

(Chapman et al. 2015; Harpur et al. 2015; Chapman et al. 2016), and further aid in differentiating Africanized bees from commercial populations. Subspecies from the other lineages have contributed to modern commercial populations (Winston 1987; Ruttner 1988), for example *A. m. caucasica* from the O lineage is particularly popular in Australia (Weatherhead 1986). SNPs specific to *A. m. scutellata* may also improve ancestry assignment given that other subspecies from Africa have been introduced to commercial populations (Sheppard 1989a, b).

The non-Africanized colony used in replicate 4 had a high level of African ancestry, and any imports from this replicate would be rejected on these grounds alone. Commercial populations from Australia, the USA and Canada typically show a minority of individuals with elevated African ancestry according to our test (Chapman et al. 2015, 2016; Harpur et al. 2015), and so this is not unusual. Other subspecies in ancient hybridization zones between Europe and Africa, such as *A. m. siciliana*, are perhaps also likely to surpass the 15% threshold; however it is unlikely that commercial beekeepers will wish to import anything other than commercial lines. Arguably, the 15% threshold is too conservative. However, this threshold was set in order to prevent all Africanized bees from being introduced to a country that does not have them, while minimizing the number of non-Africanized bees that will be rejected (Chapman et al. 2015). If the threshold were increased, there is a danger that bees of Africanized ancestry would pass the test, and this must be avoided at all costs. All three mixed semen samples from replicate 4 were found to have levels of African ancestry that would result in them being rejected for import (>15%). Another mixed semen sample, from replicate 1, would also have been rejected from import. However, the other mixed semen samples would have been accepted. All Africanized drones had high African ancestry (>56.9), while the mixed semen samples had at most 26.1% African ancestry. Thus, it is clear that the Agena

Sequenom platform is not sensitive enough to pick up alleles from a single Africanized drone, and it is not appropriate to test mixed semen samples.

Our study shows that it is not possible to detect a contribution of a small number of Africanized drones in a mixed semen sample. We therefore suggest that each drone that contributes semen is individually tested for African ancestry. Alternatively, if a queen excluder has been placed on the colony to prevent entry of foreign drones, testing of five drones will give an average proportion of African ancestry of the queen, if this is greater than 15% then drones from that colony will not pass quarantine. The UltraSEEK (Agena) platform could potentially be used to detect rare (e.g. African) alleles in mixed specimens. However, this platform is approximately ten times as expensive as the Sequenom platform (Pers. Comm. D. Hawkes, Australian Genomic Research Facility to NCC).

There are a variety of reasons to limit the movement of honeybees across the globe, particularly the possibility of disease and parasite transmission and the inadvertent establishment of invasive populations (e.g. Goulson 2003; Goka et al. 2006; Dohzono and Yokoyama 2010; Meixner et al. 2015). However, in the event that a serious pest or disease is introduced to a country, it is often desirable to introduce selected honeybee strains that tolerate the pest or disease. The risks associated with bee introductions can be minimized by quarantine procedures designed to detect pests, diseases and Africanization. In the case of Australia, we recommend the following protocol be adopted for testing semen imports for evidence of Africanization.

1. Semen should be collected in the exporting country by an approved supplier. The cadavers of all drones used to collect the semen should be collected into ethanol.
2. The semen and drones are then shipped to an approved institution in Australia. The cadavers are then analysed using the 37 SNP panel. If any males fail the threshold, the semen is destroyed. If not, the semen is released to the importer.



3. Some diseases can be transmitted via semen (Yue et al. 2006, 2007). Additional testing of the semen itself may be required, a particular consideration in areas that do not already have the disease (Roberts et al. 2015).

## ACKNOWLEDGEMENTS

This project was supported by grants from CSIRO Biosecurity Flagship Science and Innovation Award for Young People in Agriculture (NCC) and Rural Industries Research and Development Corporation Australia PRJ-00774 (BPO). The authors thank Bob Cox for sampling Africanized bees and two anonymous reviewers for their feedback on the manuscript.

## AUTHORS' CONTRIBUTIONS

NCC and BPO conceived this research and designed experiments; NCC, ALB and LB performed experiments; NCC, ALB, LB, JL, BAH, AZ, MHA, TER and BPO wrote the paper.

**Un groupe restreint de polymorphismes nucléotidiques simples pour déterminer la lignée d'origine des abeilles (*Apis mellifera*)**

**Abeilles africanisées / marqueur génétique informatif / lignée ancestrale / SNP**

**Ein gekürztes SNP Panel zur Herkunftsbestimmung von Honigbienen (*Apis mellifera*)**

**Marker zur Herkunftsbestimmung / Single Nucleotide Polymorphism / Honigbiene / *Apis mellifera* / Afrikanisierte Honigbiene**

## REFERENCES

- Alburaki M., B. Bertrand, H. Legout, S. Moulin, A. Alburaki, et al. (2013) A fifth major genetic group among honeybees revealed in Syria. *BMC Genetics* **14**: 117.
- Alqarni A.S., M.A. Hannan, A.A. Owayss, M.S. Engel. (2011) The indigenous honey bees of Saudi Arabia (Hymenoptera, Apidae, *Apis mellifera jemenitica* Ruttner): their natural history and role in beekeeping. *ZooKeys* **134**: 83–98.
- Chapman N.C., B.A. Harpur, J. Lim, T.E. Rinderer, M.H. Allsopp, et al. (2015) A SNP test to identify Africanized honeybees via proportion of 'African' ancestry. *Mol. Ecol. Res.* **15**: 1346–1355.
- Chapman N.C., B.A. Harpur, J. Lim, T.E. Rinderer, M.H. Allsopp, et al. (2016) Hybrid origins of Australian honeybees (*Apis mellifera*). *Apidologie* **47**: 26–34.
- Dietz A., R. Krell, F.A. Eischen. (1985) Preliminary investigation on the distribution of Africanized honey bees in Argentina. *Apidologie* **16**: 99–108.
- Dohzono I., J. Yokoyama. (2010) Impacts of alien bees on native plant-pollinator relationships: a review with special emphasis on plant reproduction. *Appl. Entomol. Zool.* **45**: 37–47.
- Estoup A., M. Solignac, M. Harry, J.M. Cornuet. (1993) Characterization of (GT)<sub>n</sub> and (CT)<sub>n</sub> microsatellites in two insect species: *Apis mellifera* and *Bombus terrestris*. *Nucleic Acids Res.* **21**: 1427–1431.
- Frank P., L. Garnery, M. Solignac, J.M. Cornuet. (2000) Molecular confirmation of a fourth lineage in honeybees from the Near East. *Apidologie* **31**: 167–180.
- Frank P., L. Garnery, A. Loiseau, B.P. Oldroyd, H.R. Hepburn, et al. (2001) Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data. *Heredity* **86**: 420–430.
- Goka K., K. Okabe, M. Yoneda. (2006) Worldwide migration of parasitic mites as a result of bumblebee commercialization. *Popul. Ecol.* **48**: 285–291.
- Goulson D. (2003) Effects of introduced bees on native ecosystems. *Annu. Rev. Ecol. Evol. Syst.* **34**: 1–26.
- Harpur B.A., N.C. Chapman, L. Krimus, P. Maciukiewicz, V. Sandhu, et al. (2015) Assessing patterns of admixture and ancestry in Canadian honey bees. *Insectes Soc.* **62**: 479–489.
- Ilyasov R.A., A.V. Poskryakov, A.G. Nikolenko. (2015) New SNP markers of the honeybee vitellogenin gene (*Vg*) used for identification of subspecies *Apis mellifera mellifera* L. *Russian J. Genetics* **51**: 163–168.
- Jarnevich C.S., W.E. Esaias, P.L.A. Ma, J.T. Morissette, J.E. Nickeson, et al. (2014) Regional distribution models with lack of proximate predictors: Africanized honeybees expanding north. *Divers. Distrib.* **20**: 193–201.
- Meixner M.D., P. Kryger, C. Costa. (2015) Effects of genotype, environment and their interaction on honey bee health in Europe. *Curr. Opin. Insect Sci.* **10**: 177–184.
- Muñoz I., D. Henriques, J.S. Johnston, J. Chávez-Galarza, P. Kryger, et al. (2015) Reduced SNP panels for genetic identification and introgression analysis in the dark honey bee (*Apis mellifera mellifera*). *PLoS One* **10**: e0124365.
- Oldroyd B.P., N.C. Chapman. (2016) A genetic test for Africanized honey bees, Rural Industries Research and Development Corporation, Canberra, Australia.
- Oldroyd B.P., M.H. Allsopp, J. Lim, M. Beekman. (2011) A thelytokous lineage of socially parasitic honey bees has retained heterozygosity despite at least 10 years of inbreeding. *Evolution* **65**: 860–868.

- Pinto M.A., D. Henriques, J. Chávez-Galarza, P. Kryger, L. Garnery, et al. (2014) Genetic integrity of the Dark European honey bee (*Apis mellifera mellifera*) from protected populations: a genome-wide assessment using SNPs and mtDNA sequence data. *J. Apic. Res.* **53**: 269–278.
- Pritchard J.K., M. Stephens, P. Donnelly. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Roberts J.M.K., D.L. Anderson, P. Durr. (2015) Upgrading knowledge on pathogens (particularly viruses) of Australian honey bees, Canberra.
- Ruttner F. (1988) *Biogeography and Taxonomy of Honeybees*. Springer-Verlag, Berlin.
- Schneider S.S., G. DeGrandi-Hoffman, D.R. Smith. (2004) The African honeybee: factors contributing to a successful biological invasion. *Annu. Rev. Entomol.* **49**: 351–376.
- Sheppard W.S. (1989a) A history of the introduction of honey bee races into the United States. Part I. *Am. Bee J.* **129**: 617–619.
- Sheppard W.S. (1989b) A history of the introduction of honey bee races into the United States. Part II. *Am. Bee J.* **129**: 664–667.
- Walsh P.S., D.A. Metzger, R. Higuchi. (1991) Chelex (R)100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* **10**: 507.
- Weatherhead T. (1986) *Boxes to Bar Hives: Beekeeping History of Queensland*. International Colour Productions, Stanthorpe.
- Winston M.L. (1987) *The Biology of the Honey Bee*. Harvard University Press, Cambridge.
- Yue C., M. Schröder, K. Bienefeld, E. Genersch. (2006) Detection of viral sequences in semen of honeybees (*Apis mellifera*): Evidence for vertical transmission of viruses through drones. *J. Invertebr. Pathol.* **92**: 105–108.
- Yue C., M. Schröder, S. Goisder, E. Genersch. (2007) Vertical-transmission routes for deformed wing virus of honeybees (*Apis mellifera*). *J. Gen. Virol.* **88**: 2329–2336.