Insectes Sociaux CrossMark

Assessing patterns of admixture and ancestry in Canadian honey bees

B. A. Harpur[1] · N. C. Chapman[2] · L. Krimus[1] · P. Maciukiewicz[1] · V. Sandhu[1] ·
K. Sood[1] · J. Lim[2] · T. E. Rinderer[3] · M. H. Allsopp[4] · B. P. Oldroyd[2] ·
A. Zayed[1]

Abstract Canadian honey bees, like all honey bees in the New World, originated from centuries of importation of predominately European subspecies, but their precise genetic ancestry has not been investigated. We used a citizen science approach that engaged a diverse group of beekeepers to undertake the largest population genetic study of Canadian honey bees. We used the dataset to characterize the ancestry of Canadian honey bee populations, test if Northern Canadian colonies have a greater proportion of ancestry from subspecies native to Northern Europe, and determine the effectiveness of using single nucleotide polymorphism (SNPs) to distinguish between Canadian bees and the aggressive and invasive Africanized honey bee found from South America to the Southern United States. We genotyped 855 worker honey bees at 91 ancestrally informative SNPs and found very low levels of genetic differentiation within Canada at these SNPs and small but significant differences in ancestry between provinces.

Honey bee populations in Northern and Western Canada were more closely related to subspecies from Southern and Mediterranean Europe. We attributed this pattern to differences in importation practices within Canada. Finally, we were able to accurately discriminate between Africanized bees and Canadian bees using the ancestrally informative SNPs, supporting the use of SNPs for accurately detecting Africanized honey bees and providing valuable insights into the genetic structure of Canadian bees, all while engaging beekeepers in the scientific process.

Introduction

The Western honey bee, *Apis mellifera* L., is native to the Old World where it has five major evolutionary lineages: the *A* lineage of Africa, the *C* and *M* lineages of Europe, and the *Y* and *O* lineages of North Eastern Africa, parts of the Middle East and Asia (Alqarni et al. 2011; Franck et al. 2001; Ruttner 1988; Whitfield et al. 2006). These lineages are delineated geographically, morphologically, and genetically and they include approximately 24 subspecies (Arias and Sheppard 1996; Franck et al. 2000; Garnery et al. 1992, 1993; Palmer et al. 2000; Ruttner 1988; Wallberg et al. 2014; Whitfield et al. 2006). The current honey bee populations of North America are the result of centuries of importation, chiefly from the two European lineages (*C* and *M*). Canada's *A. mellifera* populations originated from importations beginning with European settlers. The first introductions were likely from the *M* lineage (e.g. *A. m. mellifera*) (Cornuet 1986; Root 1985; Seeley 1985) likely

✉ A. Zayed
zayed@yorku.ca

1 Department of Biology, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada

2 Behaviour and Genetics of Social Insects Lab, School of Biological Sciences A12, University of Sydney, Sydney, NSW 2006, Australia

3 Honey-Bee Breeding, Genetics and Physiology Research Laboratory, USDA-ARS, 1157 Ben Hur Road, Baton Rouge, LA 70820, USA

4 ARC-Plant Protection Research Institute, Stellenbosch 7599, South Africa

followed by subsequent introductions from the *C* lineage (e.g. *A. m. ligustica* and *A. m. carnica*) and some minor introductions from the *O* lineage (e.g. *A. m. caucasica*) (Sheppard 1989a, b) and *A* lineage (e.g. *A. m. intermissa* and *A. m. lamarckii*) (Pinto et al. 2007; Root 1985; Sheppard 1989a, b). Each new introduction of a lineage or subspecies into Canada was usually an effort by beekeepers to introduce "favourable" traits, akin to Brother Adam's work in the United Kingdom to breed what he thought was the ideal honey bee for beekeeping (Adam 1983; Root 1985). Brother Adam's own "Buckfast Bee" is a mix of several subspecies from each lineage and is still bred by a small number of Canadian beekeepers today.

Remarkably, there has been no large-scale investigation of the genetic ancestry of Canadian honey bees despite a history older than the country itself (Crane 1999). We undertook a study of the genetics of Canadian honey bees using a citizen science approach to characterize their genetic ancestry and to study how geography and management practices influence their genetics. We also used the population genetic dataset to test the hypothesis that beekeepers in Northern Canada maintain honey bees more related to the Northern European subspecies *A. m. mellifera* (*M* lineage). A similar pattern has been noted in Australia, where colonies in colder regions of Tasmania have higher proportions of *A. m. mellifera* ancestry relative to colonies in warmer regions (Oldroyd et al. 1995), supporting the hypothesis that *A. m. mellifera* is locally adapted for colder climes than the more Mediterranean *C* lineage (Le Conte and Navajas 2008; Ruttner 1988). We would, therefore, expect that *M* lineage bees would be more abundant in Northern Canada.

Finally, we tested the utility of the SNP panel for discriminating between Canadian honey bees and Africanized honey bees from the United States and Brazil. Africanized honey bees can be highly aggressive and are continuously distributed from central South America to the Southern United States (Collet et al. 2006; Rinderer et al. 1991; Sheppard et al. 1991; Szalanski and Magnus 2010). Africanized honey bees are the result of an introduction of the African lineage subspecies *A. m. scutellata* into Brazil in 1956 (Kerr 1967). Controlled crosses of Brazilian commercial honey bees with imported *A. m. scutellata* were performed to test if the resulting hybrid colonies would be better suited for Brazilian beekeeping. Unfortunately, the resulting "Africanized" colonies are often highly defensive (Breed et al. 2004; Collins et al. 1982; but see: Galindo-Cardona et al. 2013), swarm frequently, and typically abscond in response to adverse conditions (Winston 1992). Current tests for detecting Africanized honey bees based on mtDNA and wing morphometrics are not reliable: mtDNA tests miss cases of paternal Africanization (Sheppard and Smith 2000) and morphological tests are unable to detect

low to medium levels of Africanization (Guzman-Novoa et al. 1994). Canadian beekeepers import hundreds of thousands of queens from the United States annually and the chance of accidental importation of Africanized honey bees is rated as moderate to high by the Canadian government (AHRA 2013). We had previously shown that an ancestrally informative SNP panel was very successful at identifying Africanized bees in commercial honey bee populations from the United States and Australia (Chapman et al. 2015b), and we wanted to examine if the same panel will be suitable for use in Canada.

Methods

Citizen science project and population sampling

From July 2013 to June 2014, we solicited beekeepers across Canada to voluntarily take part in a honey bee genotyping study. Solicitations were made through social media, our personal websites, telephone, and announcements at beekeeping meetings. Beekeepers indicated their willingness to join our study by filling in an online form. This information was used to send each beekeeper a pamphlet containing sampling instructions, a small survey, sampling tubes, and a return envelope (Supplemental Files 1 and 2). Beekeepers were instructed to sample two workers (diploid females) per colony, from up to six colonies in their operation. We asked beekeepers to identify the location of their colonies, the number of colonies they manage, and the location of their queen breeder. In total, 145 beekeepers from 9 provinces and 1 territory submitted a total of 857 sampling tubes (Fig. 1a; Supplemental Datasets 1, 2).

DNA extraction and SNP genotyping

We extracted DNA from a single diploid worker from each sampling tube returned to us ($N = 857$ individual bees). High molecular weight DNA was extracted with phenol–chloroform isoamyl alcohol (25:24:1) from half of a bee's thorax. Each sample was then purified using EMD Multiscreen Millipore purification (Merck) and genotyped using the Sequenom MassARRAY MALDI-TOF (Agena) system in four multiplexes at Génome Québec's Innovation Centre.

The SNP panel was created to differentiate between each of the three major lineages thought to be most abundant in North American honey bee populations: *C*, *M*, and *A* (Chapman et al. 2015a, b). Briefly, SNPs were randomly chosen from a set of more than 20,000 with high genetic differentiation as measured by pairwise F_{st} (Weir and Cockerham 1984) between each population group from a previous full-genome re-sequencing study (Harpur

et al. 2014) and conditioned on being at least 5000 bp apart. We included an additional 19 SNPs from a previous study that also showed high genetic differentiation between European and African honey bees (Table S3 in Whitfield et al. 2006). The final panel of 144 SNPs was chosen based on its ability to be multiplexed in an inexpensive SNP genotyping platform using the Sequenom Assay Design Suite (v1.0 Sequenom, CA, USA). In the final panel, all SNPs were separated by at least 45.9 kb (average 1734.8 kb) and were effectively unlinked because of the honey bee's very high recombination rate (19 cM/Mb; Beye et al. 2006).

Population admixture analyses

To estimate each sample's ancestry, we used STRUCTURE v2.3.4 (Pritchard et al. 2000) using all polymorphisms with minimum allele frequency >0.05 and only for markers with a call rate >0.66 ($N = 91$ markers of 144 on the panel met these criteria). Similarly, we only included samples that could be successfully genotyped at 66 % of all markers ($N = 855$ samples). We evaluated population structure using a burn-in phase of 50,000 iterations followed by 100,000 Markov-Chain Monte Carlo iterations with admixture assumed and uncorrelated allele frequencies. We included in each STRUCTURE run a set of 29 reference bees, known to be of pure descent from each of the three major lineages (African: A, Western and Northern Europe: M, and Eastern and Southern Europe: C). The reference bees were used in three previous population genetic analyses performed by our group (Harpur et al. 2012, 2014; Harpur and Zayed 2013); their genotype at each of the 91 SNPs was extracted from their full genome sequences. To reduce the influence of the large query population compared against a smaller reference population and to increase processing speed by parallelizing runs, we divided the dataset into ten smaller datasets consisting of the reference population and 85 randomly selected samples. No a priori information was provided regarding population identity or location. We performed ten replicates for each of $K = 1$ to 4 populations. We used STRUCTURE HARVESTER v0.6.94 (Earl and Vonholdt 2012) to estimate the most appropriate fit of K and to implement Evanno's method for estimating ΔK (Evanno et al. 2005). For each sample, we then identified the genomic contribution of each ancestral lineage (e.g. 70 % C, 20 % M, and 10 % A) and the level of admixture ($1 -$ maximum ancestry; e.g. if a bee is 70 % C, 20 % M, and 10 % A, then admixture $= 1 - 0.7$). Finally, we used GENEPOP v4.0.11 (Raymond and Rousset 1995) to report F_{st} statistics among provinces and countries and tested if pairwise F_{st} was significant using ARLEQUIN v3.5.12 with a False Discovery Rate at $\alpha = 0.05$ (Excoffier and Lischer 2010).

Accuracy of a SNP panel as a diagnostic test for Canada

To investigate the utility of our SNP panel to detect Africanized honey bees among Canadian imports, we determined how well our panel is able to accurately classify Africanized honey bees as being African, and Canadian honey bees as European based on their genotypes. We replicated a two-stage procedure used previously to identify a cut-off at which the proportion of African ancestry is indicative of an Africanized honey bee (Chapman et al. 2015a, b).

We first estimated the True Positive rate by identifying at what minimum proportion of African ancestry (5–60 % in 5 % increments) would samples known to be Africanized (true Africanized bees) correctly be identified as such. True Africanized samples were obtained from populations in Brazil ($N = 55$) and the United States ($N = 86$) (Chapman et al. 2015a, b). We also included $A.\ m.\ capensis$ clonal lineage ($N = 3$), $A.\ m.\ capensis$ ($N = 104$) and $A.\ m.\ scutellata \times capensis$ hybrids ($N = 17$), and 128 $A.\ m.\ scutellata$ as samples that should be correctly identified as African. At each cut-off we determined the proportion of African/Africanized samples correctly identified as African.

We next estimated the False Positive rate by repeating the above analysis to identify at what maximum proportion of African ancestry true non-Africanized bees would be incorrectly classified as African. Our true non-Africanized samples were represented by the reference C and M populations (see above) as well as commercial and feral Australian samples ($N = 93$) and commercial non-Africanized samples from Canada ($N = 10$) and the United States ($N = 55$). All reference samples were previously genotyped in an Australian study at 95/144 of the SNPs available on the panel (those having minimum allele frequency >0.05 and call rate >0.66) (Chapman et al. 2015a, b). Of the 95 SNPs in this previous study, 81 were shared with the current study (81/91 markers from $N = 855$ bees with minimum allele frequency >0.05 and call rate >0.66). Therefore, for all between-country analyses, including identifying cut-offs (above), we took only the genotypes of our Canadian samples at these 81 SNPs common between the two studies SNPs.

Statistical analyses

All statistical analyses were performed in R v3.2.0 (R Core Team 2010). For geographic analyses, we binned statistics into 0.5° latitudinal and longitudinal bins. We identified trends across provinces both between provinces and as groups. We grouped Prairie Provinces (Alberta, Saskatchewan and Manitoba) and compared admixture levels among bees from Western Provinces and Territories (Yukon and British Columbia), Ontario and Quebec, and the Maritimes

(Newfoundland, New Brunswick, and Nova Scotia). When performing multiple family wise statistical tests, we corrected for False-Discovery Rate using the Benjamini and Hochberg method (1995) at $\alpha = 0.05$. Our datasets are available on GitHub (https://github.com/harpur/CanadAdmix) and as Supplemental Material (Supplemental Dataset 1–2).

Results

Sampling overview

We sent a total of 1633 individual sampling tubes across Canada and received back 857, a 52.4 % return rate. Most samples came from British Columbia ($N = 243$) and Ontario ($N = 199$; Fig. 1a; Supplemental Dataset 2). From each returned sampling tube, we genotyped a single (diploid) worker honey bee. Only two workers could not be successfully genotyped at 66 % of all markers, so all population genetic and ancestry analyses were performed on 855 Canadian samples. Beekeepers could self-report the origins of their colonies. Of those who did self-report we found that most of the samples were bred in Canada

($N = 665$). Samples of workers from queens bred outside of Canada originated from the United States ($N = 71$), New Zealand ($N = 27$) or Denmark ($N = 2$). The beekeepers that responded to our study managed between 1 and 10,500 colonies (mean $= 400 \pm 58.9$ SE), indicating that we successfully solicited interest from both hobbyists with a few colonies and commercial beekeepers with hundreds to thousands. We asked beekeepers to self-identify the subspecies or race of their bees. We received this information for 574 out of 855 colonies sampled and genotyped in this study. The largest proportions of beekeepers identified their bees as "Italian" (30.2 %) or "Mixed" (13.7 %) (Supplemental Dataset 1).

Admixture of Canadian honey bees

Analyses using STRUCTURE significantly supported models with $K = 3$ ancestral populations (A, M and C) both with the lowest average $\text{Ln}[P(D)] = -1436.21$ method and by using Evanno's method to calculate ΔK (Supplemental Figure 1; Fig. 1b). Canadian bees were not classified as a distinct population, but instead a mix of the three ancestral lineages (Fig. 1b). Canadian samples had, on average, a
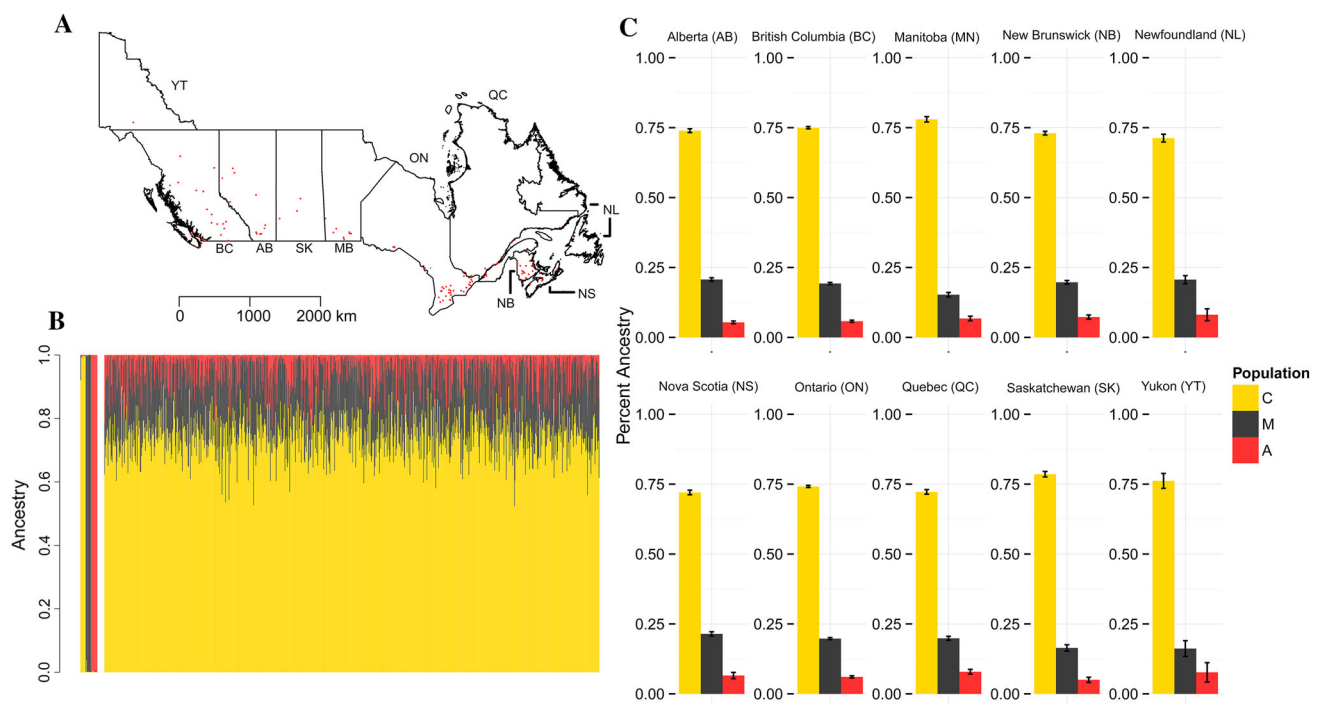

Fig. 1 a Map of sampling locations (*red dots*) and **b** average proportion ancestry in each province with province code (*yellow C*; *black M*, *red A*). Ancestry of Canadian honey bees to major honey bee lineages. The first 29 *solid bars* are known reference samples of *C*, *M*, and *A* lineage bees. All bars following the *white gap* represent 855 Canadian honey bee samples. **c** Proportion of ancestry derived from each major lineage within each sampled Canadian province. We found small but significant differences in the proportions of

C ($P = 1.9 \times 10^{-8}$) and M ancestry ($P = 3.7 \times 10^{-7}$) among provinces, but did not detect differences in the level of A ancestry ($P = 0.091$). High C (low M) ancestry is more common in the Prairie Provinces (Alberta, Saskatchewan and Manitoba) than in the Western Provinces and Territories (Yukon and British Columbia). Quebec and Ontario, and Maritime Provinces (Newfoundland, New Brunswick, and Nova Scotia), which had significantly lower C ancestry (color figure online)

large proportion of their ancestry originating from the C group (mean 74.2 %), with the remainder consisting of M group (19.6 %) and A group (6.2 %; Fig. 1b, c). As a result, differences in admixture between Canadian honey bee populations were driven by the level of M and A ancestry: increasing M and/or A ancestry lead to increased admixture (Spearman Correlation; M: $\rho = 0.56$; $P < 2.2 \times 10^{-16}$; A: $\rho = 0.51$; $P < 2.2 \times 10^{-16}$). We found small but significant differences in the level of admixture between provinces (Supplemental Figure 2; ANOVA; $F_{9,845} = 6.167$; $P = 1.9 \times 10^{-8}$). These differences tended to be between Prairie Provinces (Alberta, Saskatchewan and Manitoba) and others (Supplemental Figure 2). We confirmed this trend by pooling the bees from the Prairie Provinces and comparing their admixture to bees from Western Provinces and Territories (Yukon and British Columbia), Ontario and Quebec, and the Maritimes (Newfoundland, New Brunswick, and Nova Scotia). From this comparison, we found populations in the Prairie Provinces had lower levels of admixture when compared to populations in each of Canada's other major geographic regions (ANOVA; $F_{3,851} = 8.424$; $P = 1.6 \times 10^{-5}$; Tukey's HSD $P < 0.0035$; Supplemental Figure 2).

Provinces also differed in their patterns of ancestry. We found small but significant differences in the average mean proportion (per sample) of C (Fig. 1c; ANOVA; $F_{9,845} = 6.167$; $P = 1.9 \times 10^{-8}$) and M ancestry (ANOVA; $F_{9,845} = 5.36$; $P = 3.7 \times 10^{-7}$) among provinces, but did not detect differences in the level of A ancestry ($P = 0.091$). High C (low M) ancestry is more common in the Prairie Provinces than in the Western Provinces, Quebec and Ontario, and Maritime Provinces, which had significantly lower C ancestry (Supplemental Figure 3; ANOVA; $F_{3,851} = 8.424$; $P = 1.6 \times 10^{-5}$) and a trend towards higher levels of M ancestry (ANOVA; $F_{3,851} = 2.81$; $P = 0.0382$; Tukey's HSD $P > 0.052$). Finally, Canadian provinces had very low levels of differentiation at the loci examined (mean $F_{st} = 0.0078$; Table 1).

We found no significant evidence that samples from any self-identified subspecies or group has more A lineage ancestry than any other; however, we tested if Buckfast bees, known to have been crossed with A lineage subspecies, had higher A ancestry than other self-identified subspecies or groups. We found that Buckfast bees ($A = 8.1$ %; $N = 33$) tended to have higher levels of A ancestry than non-Buckfast bees ($A = 6.0$ %; one-tailed t test; $P = 0.06$) and the sample with the highest proportion of A ancestry within Canada ($A = 30.1$ %) is of Buckfast origin.

Distributions of honey bee lineages across Canada: local adaptation or management practices?

We predicted that Northern Canada could favour genotypes derived from honey bee subspecies accustomed to similar environments, such as the M group subspecies (Le Conte and Navajas 2008; Ruttner 1988). To test this hypothesis, we investigated associations between ancestry (C, M, or A) and geographic location (Fig. 2). Following corrections for False Discovery Rate (Benjamini and Hochberg 1995), we found a significant negative correlation between M ancestry and latitude ($P = 0.008$; $r = -0.48$) and a positive relationship between C ancestry and latitude ($P = 0.0046$; $r = 0.51$) indicating that colonies in Northern Canada tended to have higher proportions of C lineage (Fig. 2). There was no trend for A ancestry ($P = 0.45$). In addition, there was a negative correlation between admixture and latitude ($P = 0.0066$; $r = -0.14$). We found that there was a significant positive correlation between M ancestry and longitude ($P = 0.014$; $r = 0.29$), a negative relation between C ancestry and longitude ($P = 0.0006$; $r = -0.4$), and a trend for more A ancestry in Eastern Canada ($P = 0.052$; $r = 0.23$; Fig. 2).

It may be that the relationship between M lineage ancestry and geography are not a result of local adaptation but by differences in beekeeping practices. For example, small-scale beekeepers may prefer different subspecies of honey bee than commercial beekeepers or there may be regional importation differences. On the former, we found no relationship between the number of colonies managed by a beekeeper and the levels of C, M nor A ancestry of his/her samples (Spearman's rank correlation, $P > 0.38$ for all comparisons), nor level of admixture of his/her colonies and the number of colonies managed ($P = 0.46$).

We did, however, find that regional importation practices influenced ancestry. Beekeepers at latitudes $>50°$ self-reported purchasing more queens outside of Canada than beekeepers at lower latitudes ($<50°$; 22 vs 14 %; Fisher exact test; $P = 0.039$). Western beekeepers (longitude $<-100°$) also reported importing more queens than Eastern beekeepers ($>-100°$; 17.1 vs 9.4 %; $P = 0.018$). We found that imported colonies (colonies reported to have been bought outside of Canada) had significantly more C ancestry (ANOVA; $F_{1,763} = 18.21$; $P = 2.2 \times 10^{-5}$) and significantly less M ($F_{1,763} = 5.096$; $P = 0.024$), and A ancestry ($F_{1,763} = 5.82$; $P = 0.0218$; Fig. 3) than Canadian-bred and purchased colonies.

Table 1 Pairwise fixation index (F_{st}) between each Canadian province in this study

	British Columbia	Alberta	Saskatchewan	Manitoba	Ontario	Quebec	New Brunswick	Nova Scotia	Newfoundland
Alberta	0.0013								
Saskatchewan	0.0021	0.0016							
Manitoba	0.0035	0.0017	0.0026						
Ontario	0.0019	0.0019	0	0.0027					
Quebec	0.0020	0.0017	0.0042	0.0043	0.0026				
New Brunswick	0.0030	0.0047	0.0019	0.0089	0.0047	0.0043			
Nova Scotia	0.0046	0.0062	0.0104	0.0133	0.0052	0.0028	0.0077		
Newfoundland	0.0015	0.0020	0	0.0022	0.0001	0.0099	0.0045	0.0131	
Yukon Territory	0.0264	0.0253	0.0242	0.0257	0.0259	0.0248	0.0246	0.0209	0.0059

No comparisons were significant (FDR < 0.05)

Fig. 2 Relationships between latitude, longitude and percent ancestry (percentage C, M and A). Latitude is negatively correlated with M ancestry and is positively correlated with C ancestry, but was not significantly correlated to A ancestry. Longitude positively correlated with M ancestry, negatively with C ancestry but was not significantly correlated with A ancestry
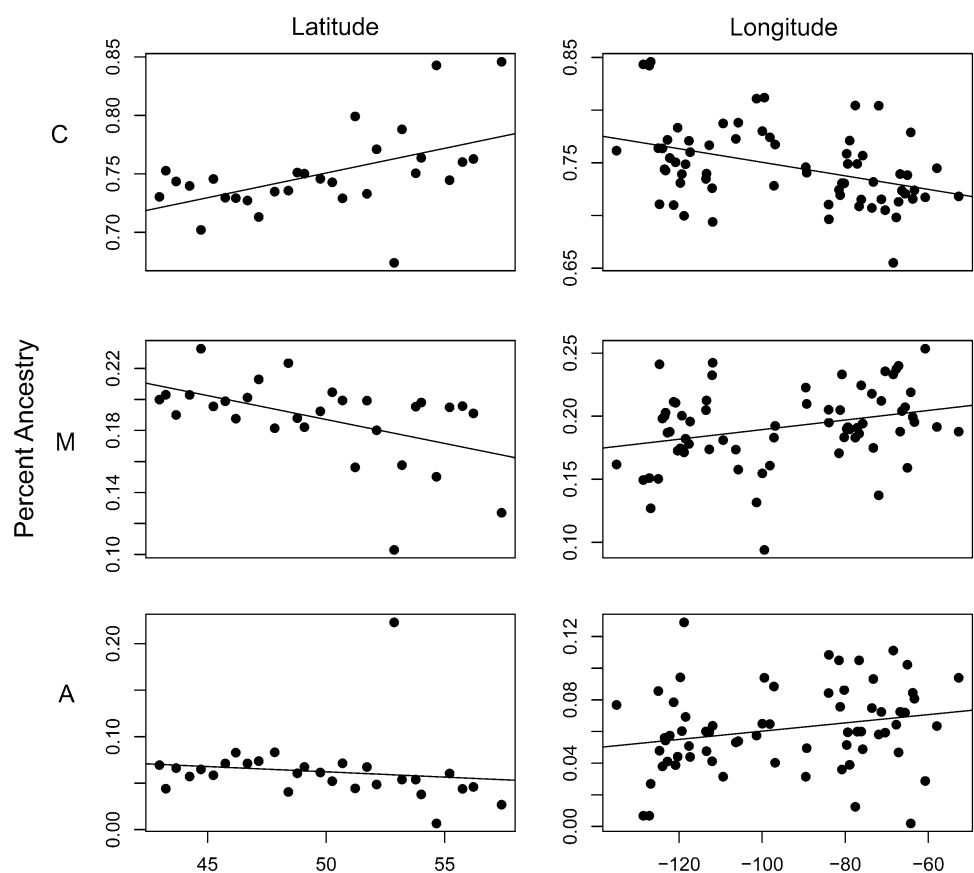
Admixture on a global scale

We compared our dataset of Canadian honey bee ancestries to those of commercial honey bee populations in Australia and non-Africanized populations in the United States that were genotyped using the same SNP panel (Chapman et al. 2015a). We found significant differences in the levels of admixture between countries (ANOVA; $F_{2,1000} = 33.1$; $P < 1.2 \times 10^{-14}$), with Canadian samples (mean = 25 %)

having similar levels of admixture as Australian samples (Tukey's HSD; $P = 0.054$; mean = 31 %) and both having higher admixture than United States commercial samples (Tukey's HSD; $P < 0.00001$; mean = 23 %). We found no differences in the level of African ancestry of commercial colonies between these countries ($P = 0.297$), but we did find significant differences in the level of M ancestry ($F_{2,1002} = 96.95$; $P < 2.2 \times 10^{-16}$) with significantly higher levels in commercial Australia (mean = 30.5 %)
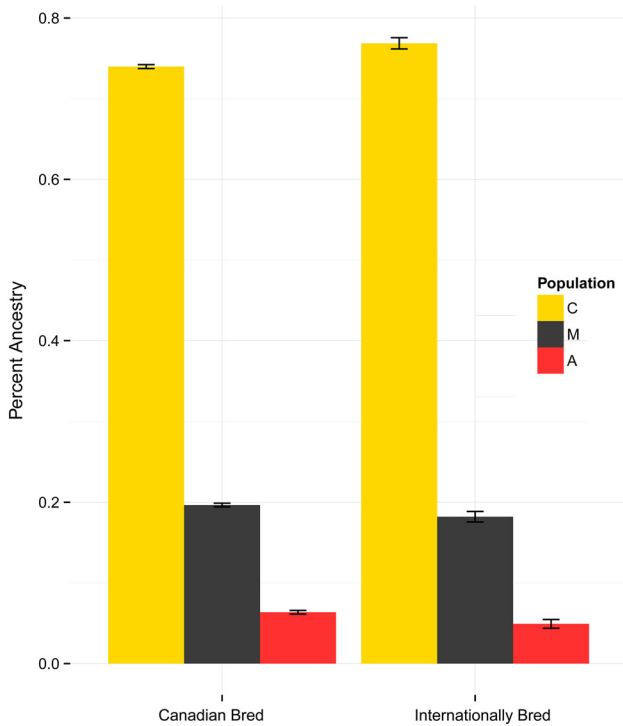
Fig. 3 Ancestry and country-of-origin of Canadian honey bee stocks. Canadian-bred colonies had significantly less C ancestry relative to internationally purchased colonies

Table 2 Pairwise fixation index (F_{st}) between Canadian, Australian and United States honey bee colonies

	Canada	Australia
Australia	0.050	
United States	0.038	0.040

No comparisons were significant (FDR < 0.05)

relative to both Canada (19.3 %) and the United States (Tukey's HSD; $P < 0.0001$; 18.0 %;). An inverse trend was found for C ancestry: Canada (74.2 %) and the United States (76.6 %) had more C ancestry than Australia (Tukey's HSD; $P < 0.0001$; 64.1 %). Even with these differences, we found low levels of differentiation between countries: average F_{st} between countries was 0.04 (Table 2).

Accuracy of the SNP based Africanization test in Canada

Africanized bees have higher levels of African ancestry (over 50 %), compared to non-Africanized bees (less than 25 %) allowing us to distinguish potentially Africanized samples using a predetermined threshold of African ancestry (Chapman et al. 2015a, b). We re-evaluated this cut-off in light of the ancestry of Canadian honey bee
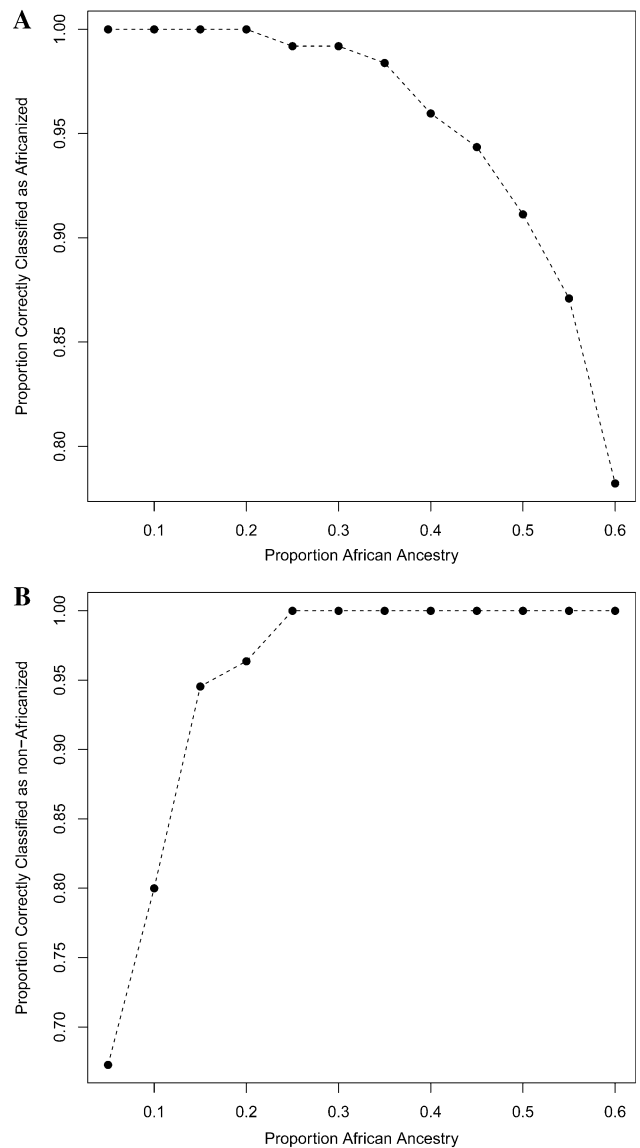

Fig. 4 a To identify an optimal true positive rate, we estimated the proportion of African ancestry at which true-Africanized ($N = 393$) bees collected from source populations in Africa, Brazil and the United States would be correctly identified as Africanized. **b** To identify an optimal false positive rate, we estimated the proportion of African ancestry at which all true-non-Africanized bees ($N = 187$) would be correctly identified as not Africanized (i.e. not incorrectly identified as African) using 5 % increments of A ancestry

populations (Fig. 4a). When we used a minimum cut-off threshold of 15 % African ancestry or greater, we obtained a True Positive rate of 1 and all true Africanized samples were correctly identified as such. When this same threshold was applied to true non-Africanized commercial stocks we obtained a False Positive rate of 0.05 (95 % of true non-Africanized commercial stocks were classified as not African). At a more conservative threshold (25 % A), we obtained a False Positive rate of 0 (100 % of true non-Africanized commercial stocks were classified as not

African; Fig. 4b). Therefore, using the more conservative cut-off threshold of 25 %, which has the maximum True Positive rate and minimum False Positive rate, we found that 99.82 % of the 855 Canadian honey bees genotyped herein could be classified as not African, as expected (Fig. 1).

Discussion

Patterns of admixture within Canada

Canada has no native populations of A. mellifera; resident populations are the result of centuries of importation predominately from the C and M lineages of Europe (e.g. A.m. ligustica and A. m. mellifera) (Cornuet 1986; Pinto et al. 2007; Seeley 1985; Sheppard 1989a, b). We have demonstrated here that contemporary Canadian honey bees are largely derived of C lineage subspecies, similar to populations in the United States (Delaney et al. 2009; Pinto et al. 2007; Seeley 1985; Sheppard 1988, 1989a, b) and Australia (Chapman et al. 2015b; Oxley and Oldroyd 2009). This pattern is likely a result of both North American and Australian beekeepers long favouring C lineage bees for their docility and higher honey production (Langstroth and Dadant 1889). Beekeepers have regularly imported and admixed local populations with A. m. ligustica (a practice once called "Italianizing") to introduce these favourable phenotypes (Jensen et al. 2005; Moritz et al. 2005). The large C lineage component of Canadian honey bees is likely a result of past importation preferences and the use of "Italianized" colonies that continues today.

Previous studies have discovered differences in ancestry between feral and commercial populations (Chapman et al. 2008, 2015a; Delaney et al. 2009; Schiff and Sheppard 1995; Sheppard 1988), with feral bees having higher levels of M ancestry. This pattern is thought to be the result of beekeepers either favouring the use of C lineage bees or selection in feral populations favouring M ancestry (Pinto et al. 2005). We did not include feral populations in our Canada survey. However, we did find that a colony's location was correlated with its ancestry. North-western Canada had more C ancestry (less M) than South-eastern Canada. This is counter to expectation as northern colonies were expected to be comprised of more northern-derived (i.e. M lineage) ancestry (Oldroyd et al. 1995). We attribute this pattern not to selective differences between parts of the country, but rather to beekeepers in North-western Canada self-reporting that they import more queens (colonies) from international sources that have higher C ancestry than colonies reported to be purchased within Canada.

Commercial honey bee populations have been noted previously for their relatively low levels of differentiation within their introduced ranges (Chapman et al. 2015a, b; Delaney et al. 2009; Harpur et al. 2012). Three factors contribute to this pattern: high gene flow, similar importation histories, and the relatively young age of commercial populations. The Canadian samples included in this study were separated by as many as 4772 km and our international samples much further. Our data suggest that gene flow within Canada is very high; most beekeepers (86.9 %) reported purchasing queens from breeders within Canada rather than rearing their own queens. Inter-population comparisons herein support that commercial colonies have very low levels of differentiation at the loci examined: we found the lowest levels of differentiation between commercial US and Canadian populations (Table 2), two populations that exchange honey bees frequently. Canadian beekeepers import 150,000–200,000 queen bees from the United States each year (Tavares 2014). Commercial populations in Canada, the United States and Australia are also relatively young and originate from similar source populations. North America has only had resident populations of honey bees since the seventeenth century (Sheppard 1989a, b) and much like Australia and the United States (Chapman et al. 2008; Hopkins 1886; Jolly 2004; Koulianos and Crozier 1996, 1997; Oldroyd et al. 1992, 1995; Oxley and Oldroyd 2009; Ruttner 1976; Sheppard 1989a, b), the Canadian populations examined herein were likely first derived from the M lineage and later shifted to C lineage. Because North American and Australian populations are relatively young, drift has less time to alter allele frequencies, and potential differences are flooded by gene flow. Taken together, the young age of these populations, their similar importation histories, and high gene flow have likely contributed to the current low levels of genetic differentiation and patterns of admixture.

Admixture in global commercial populations

While introgression can be detrimental to the conservation of honey bees within their native ranges (De la Rua et al. 2009, 2013; Meixner et al. 2010; Pinto et al. 2014), it is actively sought after in regions without native A. mellifera populations, such as North America (Cobey et al. 2012; Sheppard 2012). It has been well documented that genetic diversity is important to the health of colonies (Jones et al. 2004; Mattila and Seeley 2007; Tarpy 2003), and beekeepers seek novel genotypes resistant to pests (Cobey et al. 2012; Rinderer et al. 2010; Sheppard 2012). Admixture has been shown to increase levels of genetic diversity in honey bees (Harpur et al. 2012, 2013) and beekeepers have been intentionally interbreeding subspecies of honey bee for at least a century in North America (e.g. Root 1985). Using

tools such as the SNP panel herein (Chapman et al. 2015b) it can be possible for regulators to target and manage the introduction of novel genetic stock. A corollary is that these SNP panels can also be used to test for introgression of unwanted genetic stock such as *C* lineage in *M* lineage bees (Munoz et al. 2015; Pinto et al. 2014) or Africanized bees in North America.

The utility of a SNP based assay for monitoring Canadian imports

The current tests available to distinguish Africanized from non-Africanized colonies prior to importation can be unreliable and as such the risk of importation of Africanized honey bees into Canada has been declared moderate to high by the Canadian Food Inspection Agency (AHRA 2013). The SNP panel used in this study was designed to identify bees with African ancestry regardless of their maternal or paternal backgrounds. Using the frequency of African ancestry, we were able to demonstrate that Africanized honey bees can confidently be detected: we were able to detect true Africanized bees with 100 % accuracy with a False Positive rate less than 5 %.

We found low but pervasive levels of African ancestry in Canadian honey bees. Levels of African ancestry in Canadian bees ranged from 0.1 to 33 %, very similar to levels of African ancestry found in Australia (0.3–32.8 %; Chapman et al. 2015b). To our knowledge, there have been no deliberate introductions of Africanized bees into Canada. We instead propose that the low level of African ancestry in Canadian bees likely resulted from early importations of *A* lineage subspecies other than *A. m. scutellata*. Most likely, Canadian beekeepers imported *A. m. intermissa* (Seeley 1985), *A. m. lamarkii* (Nielsen et al. 2000), or other North African subspecies. As well, Canadian beekeepers have imported bees admixed with other African lineages, such as the Buckfast bee. Beekeepers in Ontario have maintained Buckfast bees—those developed by Brother Adam— since the 1960s (Otis, Pers. Comm.). Buckfast bees were originally crossed with the *A* lineage subspecies *A. m. saharensis* (Adam 1983). The original Buckfast bees brought into Ontario were partially derived from *A. m. saharensis* and *A. m. monticola* (Otis, Pers. Comm.). Even with the few Buckfast bees represented in our dataset ($N = 33$), we found a trend for Buckfast bees having higher *A* lineage ancestry relative to all other subspecies or groups identified by beekeepers. Although we are unable to differentiate African subspecies with the current version of the SNP panel, the addition of informative alleles for each *A* lineage subspecies, particularly *A. m. scutellata*, will enable us to better determine the origins of this pattern in the future.

Conclusions

Our data are the first in-depth assessment of the genetic structure of honey bees in Canada. Honey bees in this country, like most in the world, live predominantly in managed populations and management practices have significantly impacted their genetic structure and admixture, as we have demonstrated here and elsewhere (Harpur et al. 2012, 2013). How these management practices influence feral populations or contribute to the long-term success of managed populations still remain largely unanswered questions. Using our approach as a model and incorporating public engagement and high-throughput genetic data could help to answer these questions as well as provide valuable tools to beekeepers.

References

Adam B (1983) In search of the best strains of bees, 2nd edn. Peacock Press, Bantam

AHRA (2013) Risk assessment on the importation of honey bee (*Apis mellifera*) packages from the United States of America, Canadian Food Inspection Agency, vol 13, Canada

Alqarni AS, Hannan MA, Owayss AA, Engel MS (2011) The indigenous honey bees of Saudi Arabia (Hymenoptera, Apidae, *Apis mellifera jemenitica* Ruttner): their natural history and role in beekeeping. Zookeys. doi:10.3897/zookeys.134.1677

Arias MC, Sheppard WS (1996) Molecular phylogenetics of honey bee subspecies (*Apis mellifera* L.) inferred from mitochondrial DNA sequence. Mol Phylogenet Evol 5:557–566

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate—a practical and powerful approach to multiple testing. J R Stat Soc B Metthodol 57:289–300

Beye M, Gattermeier I, Hasselmann M, Gempe T, Schioett M, Baines JF, Schlipalius D, Mougel F, Emore C, Rueppell O, Sirvio A, Guzman-Novoa E, Hunt G, Solignac M, Page RE (2006) Exceptionally high levels of recombination across the honey bee genome. Genome Res 16:1339–1344. doi:10.1101/Gr.5680406

Breed MD, Guzman-Novoa E, Hunt GJ (2004) Defensive behavior of honey bees: organization, genetics and comparisons with other bees. Annu Rev Entomol 49:271–298

Chapman NC, Lim J, Oldroyd BP (2008) Population genetics of commercial and feral honey bees in Western Australia. J Econ Entomol 101:272–277

Chapman NC, Harpur BA, Lim J, Rinderer TE, Allsopp MH, Zayed A, Oldroyd BP (2015a) Hybrid origins of Australian honey bees (*Apis mellifera*). Apidologie. doi:10.1007/s13592-015-0371-0

Chapman NC, Harpur BA, Lim J, Rinderer TE, Allsopp MH, Zayed A, Oldroyd BP (2015b) A SNP test to identify Africanized honeybees via proportion of "African" ancestry. Mol Ecol Resour. doi:10.1111/1755-0998.12411

Cobey S, Sheppard WS, Tarpy DR (2012) Status of breeding practices and genetic diversity in domestic US honey bees. In: Sammataro D, Yoder J (eds) Honey bee colony health: challenges and sustainable solutions. CRC Press, Boca Raton, pp 39–49

Collet T, Ferreira KM, Arias MC, Soares AEE, Del Lama MA (2006) Genetic structure of Africanized honeybee populations (*Apis mellifera* L.) from Brazil and Uruguay viewed through mitochondrial DNA COI–COII patterns. Heredity 97:329–335. doi:10.1038/sj.hdy.6800875

Collins AM, Rinderer TE, Harbo JB, Bolten AB (1982) Colony defense by Africanized and European honey bees. Science 218:72–74

Core Team R (2010) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Cornuet J-M (1986) Population genetics. In: Rinderer TE (ed) Bee genetics and breeding. Academic Press, Orlando, pp 235–254

Crane E (1999) The world history of beekeeping and honey hunting. Routledge, New York

De la Rua P, Jaffe R, Dall'Olio R, Munzos I, Serrana J (2009) Biodiversity, conservation and current threats to European honeybees. Apidologie 40:263–284

De la Rua P, Jaffe R, Munoz I, Serrano J, Moritz RFA, Kraus FB (2013) Conserving genetic diversity in the honeybee: comments on Harpur et al. (2012). Mol Ecol 22:3208–3210. doi:10.1111/mec.12333

Delaney DA, Meixner MD, Schiff NM, Sheppard WS (2009) Genetic characterization of commercial honey bee (Hymenoptera: Apidae) populations in the United States by using mitochondrial and microsatellite markers. Ann Entomol Soc Am 102:666–673. doi:10.1603/008.102.0411

Earl DA, Vonholdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361. doi:10.1007/s12686-011-9548-7

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620. doi:10.1111/j.1365-294X.2005.02553.x

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567. doi:10.1111/j.1755-0998.2010.02847.x

Franck P, Garnery L, Solignac M, Cornuet JM (2000) Molecular confirmation of a fourth lineage in honeybees from the Near East. Apidologie 31:167–180

Franck P, Garnery L, Loiseau A, Oldroyd BP, Hepburn HR, Solignac M, Cornuet JM (2001) Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data. Heredity 86:420–430. doi:10.1046/j.1365-2540.2001.00842.x

Galindo-Cardona A, Acevedo-Gonzalez JP, Rivera-Marchand B, Giray T (2013) Genetic structure of the gentle Africanized honey bee population (gAHB) in Puerto Rico. BMC Genet. doi:10.1186/1471-2156-14-65

Garnery L, Cornuet JM, Solignac M (1992) Evolutionary history of the honey bee *Apis mellifera* inferred from mitochondrial DNA analysis. Mol Ecol 1:145–154. doi:10.1111/j.1365-294X.1992.tb00170.x

Garnery L, Solignac M, Celebrano G, Cornuet JM (1993) A simple test using restricted PCR amplified mitochondrial DNA to study the genetic structure of *Apis mellifera* L. Experientia 49:1016–1021. doi:10.1007/bf02125651

Guzman-Novoa E, Page RE, Fondrk MK (1994) Morphometric techniques do not detect intermediate and low-levels of Africanization in honey bee (Hymenoptera, Apidae) colonies. Ann Entomol Soc Am 87:507–515

Harpur BA, Zayed A (2013) Accelerated evolution of innate immunity proteins in social insects: adaptive evolution or relaxed constraint? Mol Biol Evol 30:1665–1674. doi:10.1093/molbev/mst061

Harpur BA, Minaei S, Kent CF, Zayed A (2012) Management increases genetic diversity of honey bees via admixture. Mol Ecol 21:4414–4421. doi:10.1111/j.1365-294X.2012.05614.x

Harpur BA, Minaei S, Kent CF, Zayed A (2013) Admixture increases diversity in managed honey bees: reply to De la Rua et al. (2013). Mol Ecol 22:3211–3215. doi:10.1111/Mec.12332

Harpur BA, Kent CF, Molodtsova D, Lebon JMD, Alqarni AS, Owayss AA, Zayed A (2014) Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. Proc Natl Acad Sci USA 111:2614–2619. doi:10.1073/pnas.1315506111

Hopkins I (1886) Illustrated Australasian bee manual and complete guide to modern bee culture in the Southern Hemispere, 3rd edn. Isaac Hopkins, Auckland

Jensen AB, Palmer KA, Boomsma JJ, Pedersen BV (2005) Varying degrees of *Apis mellifera ligustica* introgression in protected populations of the black honeybee, *Apis mellifera mellifera*, in Northwest Europe. Mol Ecol 14:93–106. doi:10.1111/j.1365-294X.2004.02399.x

Jolly B (2004) South Australia's early ligurian beekeeping—and a lingering Kangaroo Island fable. J Hist Soc S Aust 32:69–81

Jones JC, Myerscough MR, Graham S, Oldroyd BP (2004) Honey bee nest thermoregulation: diversity promotes stability. Science 305:402–404. doi:10.1126/science.1096340

Kerr WE (1967) The history of the introduction of Africanized honey bees to Brazil. S Afr Bee J 39:3–5

Koulianos S, Crozier R (1996) Mitochondrial DNA sequence data provides further evidence that the honeybees of Kangaroo Island, Australia are of hybrid origin. Apidologie 27:165–174

Koulianos S, Crozier R (1997) Mitochondrial sequence characterisation of Australian commercial and feral honeybee strains, *Apis mellifera* L. (Hymenoptera: Apidae), in the context of the species worldwide. Aust J Entomol 36:359–363

Langstroth L, Dadant C (1889) Langstroth on the hive and honey bee. C. Dadant & Son, Hamilton

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

Mattila HR, Seeley TD (2007) Genetic diversity in honey bee colonies enhances productivity and fitness. Science 317:362–364. doi:10.1126/science.1143046

Meixner MD, Costa C, Kryger P, Hatjina F, Bouga M, Ivanova E, Buchler R (2010) Conserving diversity and vitality for honey bee breeding. J Apicult Res 49:85–92. doi:10.3896/ibra.1.49.1.12

Moritz RFA, Hartel S, Neumann P (2005) Global invasions of the western honeybee (*Apis mellifera*) and the consequences for biodiversity. Ecoscience 12:289–301. doi:10.2980/i1195-6860-12-3-289.1

Munoz I, Henriques D, Johnston JS, Chavez-Galarza J, Kryger P, Pinto MA (2015) Reduced SNP panels for genetic identification and introgression analysis in the dark honey bee (*Apis mellifera mellifera*). PLoS One. doi:10.1371/journal.pone.0124365

Nielsen DI, Ebert PR, Page RE, Hunt GJ, Guzman-Novoa E (2000) Improved polymerase chain reaction-based mitochondrial genotype assay for identification of the africanized honey bee (Hymenoptera: Apidae). Ann Entomol Soc Am 93:1–6. doi:10.1603/0013-8746(2000)093[0001:Ipcrbm]2.0.Co;2

Oldroyd BP, Sheppard WS, Stelzer JA (1992) Genetic characterization of the bees of Kangaroo Island, South Australia. J Apicult Res 31:141–148

Oldroyd BP, Cornuet JM, Rowe D, Rinderer TE, Crozier RH (1995) Racial admixture of *Apis mellifera* in Tasmania, Australia—similarities and differences with natural hybrid zones in Europe. Heredity 74:315–325. doi:10.1038/hdy.1995.46

Oxley PR, Oldroyd BP (2009) Mitochondrial sequencing reveals five separate origins of 'black' *Apis mellifera* (Hymenoptera: Apidae) in Eastern Australian commercial colonies. J Econ Entomol 102:480–484

Palmer MR, Smith DR, Kaftanoglu O (2000) Turkish honeybees: genetic variation and evidence for a fourth lineage of *Apis mellifera* mtDNA. J Hered 91:42–46

Pinto MA, Rubink WL, Patton JC, Coulson RN, Johnston JS (2005) Africanization in the United States: replacement of feral European honeybees (*Apis mellifera* L.) by an African hybrid swarm. Genetics 170:1653–1665. doi:10.1534/genetics.104.035030

Pinto MA, Sheppard WS, Johnston JS, Rubink WL, Coulson RN, Schiff NM, Kandemir I, Patton JC (2007) Honey bees (Hymenoptera: Apidae) of African origin exist in non-Africanized areas of the Southern United States: evidence from mitochondrial DNA. Ann Entomol Soc Am 100:289–295

Pinto MA, Henriques D, Chaves-Galarza J, Kryger P, Garnery L, Van der Zee M, Dahle B, Soland-Reckeweg G, de la Rua P, Dall'Olio R, Carreck NL, Johnson JS (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 Apicult Res 53:269–278

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Raymond M, Rousset F (1995) Genepop (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Rinderer TE, Stelzer JA, Oldroyd BP, Buco SM, Rubink WL (1991) Hybridization between European and Africanized honey bees in the neotropical Yucatan Peninsula. Science 253:309–311. doi:10.1126/science.253.5017.309

Rinderer TE, Harris JW, Hunt GJ, de Guzman LI (2010) Breeding for resistance to *Varroa destructor* in North America. Apidologie 41:409–424. doi:10.1051/apido/2010015

Root AI (1985) ABC and XYZ of bee culture, 41st edn. A I Root Co, Medina

Ruttner F (1976) Isolated populations of honeybees in Australia. J Apicult Res 15:97–104

Ruttner F (1988) Biogeography and taxonomy of honeybees. Springer, Berlin

Schiff NM, Sheppard WS (1995) Genetic analysis of commercial honey bees (Hymenoptera, Apidae) from the Southeastern United States. J Econ Entomol 88:1216–1220

Seeley TD (1985) Honey bee ecology: a study of adaptation in social life. Princeton University Press, Princeton

Sheppard WS (1988) Comparative study of enzyme polymorphism in United States and European honey bee (Hymenoptera, Apidae) populations. Ann Entomol Soc Am 81:886–889

Sheppard WS (1989a) A history of the introduction of honey bee races into the United States. Part I. Am Bee J 129:617–619

Sheppard WS (1989b) A history of the introduction of honey bee races into the United States. Part II. Am Bee J 129:664–667

Sheppard WS (2012) Managed pollinator CAP coordinated agricultural project a national research and extension initiative to reverse pollinator decline honey bee genetic diversity and breeding towards the reintroduction of European germ plasm. Am Bee J 152:155–158

Sheppard WS, Smith DR (2000) Identification of African-derived bees in the Americas: a survey of methods. Ann Entomol Soc Am 93:159–176

Sheppard WS, Soares AEE, Dejong D, Shimanuki H (1991) Hybrid status of honey bee populations near the historic origin of Africanization in Brazil. Apidologie 22:643–652. doi:10.1051/apido:19910607

Szalanski AL, Magnus RM (2010) Mitochondrial DNA characterization of Africanized honey bee (*Apis mellifera* L.) populations from the USA. J Apicult Res 49:177–185. doi:10.3896/Ibra.1.49.2.06

Tarpy DR (2003) Genetic diversity within honeybee colonies prevents severe infections and promotes colony growth. Proc R Soc Lond B Biol Sci 270:99–103. doi:10.1098/rspb.2002.2199

Tavares A (2014) Statistical overview of the Canadian honey industry 2013. Government of Canada, Canada

Wallberg A, Han F, Wellhagen G, Dahle B, Kawata M, Haddad N, Simoes ZLP, Allsopp MH, Kandemir I, de la Rua P, Pirk CW, Webster MT (2014) A worldwide survey of genome sequence variation provides insight into the evolutionary history of the honeybee *Apis mellifera*. Nat Genet 46:1081–1088

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370

Whitfield CW, Behura SK, Berlocher SH, Clark AG, Johnston JS, Sheppard WS, Smith DR, Suarez AV, Weaver DB, Tsutsui ND (2006) Thrice out of Africa: ancient and recent expansions of the honey bee, *Apis mellifera*. Science 314:642–645

Winston ML (1992) The biology and management of Africanized honey bees. Annu Rev Entomol 37:173–193