UNIVERSITY OF CALIFORNIA, SAN DIEGO UCSD

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

Daniela Zárate
Ph.D. Candidate
ECOLOGY, BEHAVIOR AND EVOLUTION
TELEPHONE: (413) 652-8644
EMAIL: dazarate@ucsd.edu

DIVISION OF BIOLOGY
9500 GILMAN DRIVE
MAIL CODE 0116
LA JOLLA, CALIFORNIA

02/01/2021

Editor,
Molecular Ecology

Attached please find our manuscript *"Admixture in Africanized honey bees (Apis mellifera) from Panamá to San Diego, California (U.S.A.)" for consideration of publication as an original article. In it, we assess the nuclear genomic admixture, mitochondrial origins, and genetic diversity measures of honey bees from four populations stretching from the isthmus of Panamá to San Diego, California.* The Africanized honey bee (AHB) is the product of human-mediated hybridization between recognized ancestral lineages of honey bees drawn from western and eastern Europe, the Middle East, and the African continent. *The AHB represents a hybrid form that has displaced pre-existing, purely European, honey bees everywhere in the New World from its site of origination in Brazil to the current limits of its range in Argentina and California. Our study is unique in assessing genomic characteristics of several populations previously unstudied by NGS methods and in assessing the significant contribution of the Middle Eastern honey bee lineage to current honey bee populations. Our paper should be of interest to a wide range of evolutionary ecologists interested in hybridization as a creative force in evolution and in the genetic basis of success of invasive species. Additionally, Africanized honey bees are a species of significant economic and public interests due to its role as a commercial pollinator and concern over its fearsome nest defense behavior which has caused the popular press to brand it the "killer bee".*

thank you for your consideration,
Daniela Zarate

# *Admixture in Africanized honey bees (Apis mellifera) from Panamá to San Diego, California (U.S.A.)*

Daniela Zárate[1†], Thiago G. Lima[2], Jude D. Poole[3], Erin Calfee[4], Ronald S. Burton[2], Joshua R. Kohn[1]

[1] Section of Ecology, Behavior, and Evolution, Division of Biological Sciences, University of California, San Diego, La Jolla, CA
[2] Scripps Institute of Oceanography, University of California San Diego, La Jolla, CA
[3] Division of Biological Sciences, University of California, San Diego, La Jolla, CA
[4] Department of Evolution and Ecology and Center for Population Biology, University of California, Davis
[†]Corresponding author.
EMAIL: dazarate@ucsd.edu

## ABSTRACT

The Africanized honey bee (AHB) is a New World amalgamation of several subspecies of the western honey bee (*Apis mellifera*), a diverse taxon grouped into four major biogeographic lineages: A (African), M (western European), C (eastern European), and O (Middle Eastern). In 1956, accidental release of experimentally bred "Africanized" hybrids from a research apiary in Sao Paulo, Brazil initiated a hybrid species expansion that now extends from northern Argentina to northern California (U.S.A.). Here, we assess nuclear admixture and mitochondrial ancestry in 15 bees from each of four regions across this expansive range: the Isthmus of Panamá; Guanacaste, Costa Rica, Tapachula, Mexico; and San Diego, U.S.A to assess ancestry of AHB several decades following initial introduction and test the prediction that African ancestry decreases with increasing latitude. We find that AHB nuclear genomes from Central America and Mexico have majority African ancestry (Mexico, 79%; Costa Rica 90%; and Panamá 94%) with varying contributions from western and eastern European lineages. AHB from San Diego (CA) show markedly lower African ancestry (40%) with substantial genomic contributions from all four major honey bee lineages. The mitochondria of all bees sampled in Costa Rica and Panamá originated in Africa. The majority (11) of bees sampled in Mexico carried African mitochondria with the remainder carrying eastern European mitochondria. In the San Diego population, mitochondria from all four lineages are present. Genetic diversity measures from all New World populations are similar and exceed those of ancestral forms. The unique genetic makeup of the San Diego honey bee population makes it a rich source of genetic material for honey bee breeding.



***Keywords***

Africanized honey bees, *Apis mellifera*, admixture, genetic diversity, hybridization

## INTRODUCTION

Hybridization, the interbreeding of distinct genetic lineages, has long complicated taxonomic boundaries and challenged the perception of species as discrete taxonomic and evolutionary units. Many early evolutionary biologists considered hybridization an infrequent and abnormal event of limited evolutionary importance, resulting from the breakdown of natural isolating mechanisms (Dobzhansky, 1936; Mayr, 1942; reviewed in Barton, 2001). This was an attitude largely espoused by animal researchers while plant biologists, conscious of the high frequency of hybridization leading to viable offspring in plants, viewed introgression as a creator of genetic novelty upon which selection could act (Stebbins, 1950; Grant, 1981; Suarez-Gonzalez, Lexer, & Cronk, 2018).

Advances in sequencing technology and ancestry estimation have facilitated the identification of introgression, exposing heretofore undiscovered hybridization with unexpected high frequency. The recombination of distinct genetic lineages generates novel mosaic genomes on which selection can act (Anderson & Stebbins, 1954; Hedrick, 2013). At times, advanced generation admixture can be detrimental, resulting in hybrid breakdown with fitness loss due to nuclear-nuclear or nuclear-mitochondrial incompatibilities (Dobzhansky, 1936; Muller, 1942; Burton & Barreto, 2012). However, hybridization and recombination also have the potential to produce evolutionary novelty and may lead to the creation of new evolutionary units (reviewed in Dittrich-Reed & Fitzpatrick, 2013). Hybridization in the animal kingdom is now recognized as a

common and creative evolutionary force in processes of adaptation and diversification (Barton, 2001; Abbott *et al.,* 2013).

Admixture as a driver of adaptation and diversification has been widely studied across diverse taxonomic groups. Sunflower (*Helianthus)* hybrids can colonize and thrive in habitats neither parental species can occupy (Rieseberg *et al.,* 2003; Whitney *et al.,* 2015). Admixture jumpstarted the spectacular diversification and adaptive radiation recognized in African cichlids (Cichlidae) (Seehausen, 2004). Admixture between extinct Denisovans and ancient *Homo Sapiens* facilitated the transfer of high-altitude adaptation genes found in contemporary Tibetan peoples (Huerta-Sanchez *et al.,* 2014), an adaptation apparently paralleled in admixture between highland wolves and domesticated highland dogs (e.g. Tibetan mastiffs) (VonHoldt, Fan, Vecchyo, & Wayne, 2017). Of particular interest here, are the evolutionary dynamics that emerge from human-mediated hybridization (HMH), a phenomenon of increasing frequency resulting from either accidental or intentional introductions of biota to geographical areas beyond their native ranges (reviewed in Grabenstein & Taylor, 2018).

The Africanized honey bee (AHB) is one of the most well-documented examples of human-mediated hybridization. The western honey bee (*Apis mellifera*), well-known for its critical role as a pollinator in commercial agriculture, was first introduced from Europe to the American continents in the early 1500s. *Apis mellifera* is a diverse taxon, comprised of over thirty recognized subspecies clustering into four major lineages based on genetic, geographic, and morphometric data: A (African), M (western European), C (eastern European), and O (Middle Eastern) (Ruttner, 1988). Substantial variation in behavior, morphology, and genetics exists

across subspecies, even within the overarching clades. The eastern European subspecies are particularly favored in modern beekeeping due to their gentle nature and high fecundity while the African subspecies *A. m. scutellata* is disfavored due to the intensity of its nest defense behavior and propensity to abscond (abandon the nest en masse and move to another) (Ruttner, 1988).

Early honey bee importations were largely western European (M) and eastern European (C) in origin; the former dominating the 16th to 18th century introductions while the latter dominated later introductions (reviewed in Schneider, DeGrandi-Hoffman, & Smith, 2004). Honey bees from these clades are often generalized as European honey bees (EHB). Eastern European (Clade C) honey bees are now the variety of choice in commercial agriculture in the U.S.A., where pollination services of honey bees are valued at an estimated $14.5 billion (Morse & Calderone, 2000). Middle Eastern honey bees (Clade O) were introduced to the United States in the late 1880s and 1890s in much more limited quantities and their importation was phased out by the end of the 19$^{th}$ century in favor of other subspecies (Sheppard, 1989). Surprisingly, mitochondria of Middle Eastern origin continue to persist in the feral honey bee gene pool in the U.S.A. (Magnus & Szalanski, 2010; Kono & Kohn, 2015). African (A) subspecies were largely excluded from importation with the exception of the Egyptian subspecies *A. m. lamarkii* which was introduced to North America at low frequency (Schiff & Sheppard, 1993).

In 1956, 47 queens of the African subspecies (*A. m. scutellata*) were imported to Sao Paulo, Brazil for experimental breeding in an effort to create a honey bee better adapted to the tropical conditions (reviewed in Schneider *et al*., 2004). Researchers hoped to forge a honey bee that combined the tropical hardiness of *A. m. scutellata* with the honey production capabilities and

gentleness of the popular European subspecies. Admixed "Africanized" honey bees (AHB) were accidently released from the experimental apiary and spread into the surrounding countryside (reviewed in Schneider *et al*., 2004). Their subsequent expansion across the American continents over the past 60+ years is considered one of the "most spectacular biological invasions of all time" (Pinto, Rubink, Patton, Coulsen, & Johnston, 2005).

Africanized honey bees spread across South and Central America, hybridizing with and displacing pre-existing populations of European honey bees, resulting in a rapid replacement of European ancestry by African ancestry in honey bee populations (Lobo, Del Lama, & Mestriner, 1989; Smith, Taylor, & Brown, 1989; Hall & McMichael, 2001; reviewed in Schneider *et al.,* 2004). AHBs reached their southern range limit in Argentina in the 1970s at approximately 34° south latitude; presumably stopped from advancing further by the colder climate (Taylor & Spivak, 1984). The AHB reached Panamá by 1982, Costa Rica by 1986, Mexico by 1989, Texas by 1990, and California by 1994 (Kim & Oguro, 1999). Currently, honey bees with African mitochondria have been reported as far north as Sacramento and Solana counties in northern California, albeit at low frequencies (Kono & Kohn, 2015; Lin, McBroome, Rehman, & Johnson, 2017), while genomic African ancestry is likewise low or absent in bees from these higher latitudes (Calfee, Agra, Palacio, Ramirez, & Coop, 2020).

Replacement by AHB of pre-existing feral populations of European origin suggests strong ecological advantages for this hybrid form in the habitats it now occupies. Many aspects of the AHB's behavior are largely consistent with that of its African ancestor and may contribute to its ecological advantage. Generally, AHB exhibit greater reproductive rates than EHB; converting

pollen into brood at more rapid rates and dedicating more comb area to brood rearing (McNally & Schneider, 1992a, 1992b, 1996; reviewed in Fewell & Bertram, 2002). In the Neotropics, AHB colonies can increase 16-fold per year compared to 3- to 6-fold by EHB colonies (Otis 1991). In addition, AHB exhibit lower susceptibility to *Varroa* mite infestation of both brood and workers (Guzman-Novoa, Sanchez, Page Jr., & Garcia, 1996; Fewall & Bertram, 2002). Of particular public safety concern is the high degree of AHB nest defense. The AHB deposits more stings on a target, responds faster and in greater numbers, and pursues any perceived threat further than European forms (Collins, Rinderer, Harbo, & Bolten, 1982; DeGrandi-Hoffman, Collins, Martin, Shcmidt, & Spangler, 1997;). This elevated level of nest defense has motivated the popular press to brand the AHB as the "killer bee". (Winston, 1992; reviewed in Breed, Guzmán-Novoa, & Hunt, 2004).

The genomes of AHB are thought to be predominantly African in origin; except at the northern and southern range limits in Argentina and California, where the proportion of African ancestry decreases (Whitfield *et al.,* 2006; Nelson *et al.,* 2017; Calfee *et al.,* 2020). Early studies assessing Africanization used either mitochondrial or limited numbers of nuclear markers to assess admixture (Del Lama, Lobo, Soares, & Del Lama, 1990; Schiff, Sheppard, Loper, & Shimanuki, 1994; Quezada-Euan & Hinsull, 1995). The level of African ancestry in honey bee populations has been assessed by next generation sequencing (NGS) analyses utilizing thousands of single nucleotide polymorphisms (SNPs) only in Brazil, Argentina, and the southern and western U.S.A. (Whitfield *et al.,* 2006; Nelson, Wallberg, Simoes, Lawson, & Webster, 2017; Cridland *et al.,* 2017; Calfee *et al,* 2020).

Similar to NGS assessments from Brazilian populations, feral honey bees in Texas and Arizona reached high levels (63 – 75%) of African ancestry in just a few years after arrival of the first AHB (Rubink, Luévano-Martinez, Sugden, Wilson, & Collins, 1996; Pinto *et al.,* 2005; Rabe *et al.,* 2005; Whitfield *et al.,* 2006; Rangel *et al.,* 2016; Bozek *et al.,* 2018). In Southern California, Africanization is widespread in feral honey bees, although African genomic content is reported to be lower (30-40%) (Kono & Kohn, 2015; Cridland *et al*., 2017; Lin *et al*., 2017; Calfee *et al.,* 2020.

To date there are no whole-genome ancestry estimates for AHB between Brazil and the southern United States. In addition, previous genomic studies have usually not assessed the contribution of the Middle Eastern (O) lineage to the genomes of AHB (but see Whitfield *et al,*. 2006), despite the fact that the O mitochondrial type is known to persist in the population of at least some feral bee populations including Southern California (Kono & Kohn, 2015; Magnus & Szalanski, 2010). Whether and how AHB ancestry varies across geographic space (and time since hybridization), requires a comprehensive examination of AHB admixture throughout its geographic range and would shed light on admixture dynamics in a human-mediated invasive expansion event of great breadth. Here we sequence 60 whole genomes of Africanized honey bees collected from four regions separated across a distance of ~6,000 km: the isthmus of Panamá; Guanacaste NP, Costa Rica; Chiapas, Mexico; and San Diego County, CA, U.S.A. To our knowledge this is the first time that contributions from the four major clades of the western honey bee have been estimated for both nuclear and mitochondrial genomes in AHBs across a broad geographic range. The varied sampled sites offer an interesting temporal dimension as each site also reflects a distinct time since initial contact between resident European and advancing

Africanized forms. The AHB offers a unique opportunity to study a massive hybrid invasion and the patterns of genomic admixture and genetic diversity that emerge across both space and time.

## MATERIALS AND METHODS

### Sample Collection

We collected western honey bees (n = 15/country) from sites in each of four countries: Panamá; Guanacaste National Park, Costa Rica; Chiapas, Mexico; San Diego county, California, U.S.A. (Table 1). All samples were collected in June 2015 – August 2016 by hand-netting. Honey bees in Panamá were collected with an insect net while they foraged either on natural vegetation in rural areas, or on street vendor syrup dispensers in urban areas. Honey bees were collected across the isthmus of Panamá from five sites, each separated by > 5 km: Panamá City, Gamboa, Barro Colorado Island (BCI), Santa Rita Arriba, and Cólon. Individuals from Costa Rica were collected from the Santa Rosa sector of Guanacaste National Park in northwestern Costa Rica. These bees were collected from a localized region and likely originate from a small number of feral colonies. Honey bees from Mexico were collected from an apiary in the southern state of Chiapas, with each bee collected from a different hive. Honey bees from San Diego County, California, U.S.A. were workers collected while foraging on flowers. San Diego bees were collected across 15 sites each separated by > 5 km so that each likely represents a worker from a different colony. The furthest collection sites were separated by 65 km. Collection sites ranged from urban to rural settings. Due to the presence of hobbyist and agricultural beekeeping we do not rule out the

183 possibility that the captured honey bees were from managed rather than feral hives. However,
184 most honey bee foragers in San Diego are from feral hives (Kono & Kohn, 2015, and see results).
185
186 **Reference Honey Bee Genomes**
187 Reference honey bee genomes were obtained by downloading genomes sequenced by Wallberg
188 *et al*., (2014) and deposited in the NCBI (Project ID: PRJNA236426). Reference genomes were
189 generated by whole genome sequencing on a SOLiD 5500xl platform to produce 75-bp reads
190 with an average coverage of 4.4X ±1.5X per individual (Wallberg *et al.,* 2014 supplementary
191 material). For the African (A) clade, we downloaded 10 genomes of the subspecies *Apis mellifera*
192 *scutellata*, the sub-Saharan subspecies imported to Brazil. For the western European (M) clade,
193 we downloaded 20 genomes: *Apis mellifera mellifera* (n = 10) and *Apis mellifera iberiensis* (n =
194 10). For the eastern European clade (C) we downloaded 20 genomes: *Apis mellifera carnica* (n =
195 10) and *Apis mellifera ligustica* (n = 10). For the Middle Eastern (O) clade, we downloaded 20
196 genomes: *Apis mellifera anatoliaca* (n = 10) and *Apis mellifera syriaca* (n = 10) (see Table 1,
197 S1). In total we used a panel of 70 reference honey bee genomes representing the four major
198 honey bee clades and spanning 7 subspecies.
199
200 **DNA Extraction & Sequencing**
201 We extracted DNA from crushed heads of sampled honey bees using the standard protocol of the
202 Qiagen DNAeasy Blood & Tissue extraction kit. DNA purity and appropriate concentration for
203 sequencing were validated with a Qubit fluorometer prior to submission for library preparation.
204 The DNA was submitted for DNA KAPA library construction and whole-genome sequencing at
205 the Institute for Genomic Medicine (IGM) at UC San Diego. Individuals were multiplexed and

206 sequenced across three lanes of an Illumina HiSeq4000 platform using 100-bp paired end reads.
207 Average genomic coverage per individual was 29X ±1.2x.
208
209 **Sequence Filtering & Alignment**
210 Raw reads generated from sequencing, and those downloaded from NCBI, were trimmed and
211 filtered for quality and length using a PoPoolation (Kofler *et al*., 2011) perl script (trim-fastq.pl)
212 (settings: –fastq-type sanger --quality-threshold 25 --min-length 40). Filtered reads were aligned
213 to the Amel_4.5 reference genome assembled by The Honey Bee Genome Sequencing
214 Consortium (2006) using the BWA v0.7.12 bwa mem algorithm under default settings (Li &
215 Durban, 2009). Reads were then sorted, merged, and filtered again for mapping quality (quality
216 score < 20 were discarded) using Samtools (Li, 2011).
217
218 **Variant Calling and Genotype Likelihood Estimation**
219 We used the program ANGSD v0.930 (Kornliussen *et al.,* 2014) to call variant sites and estimate
220 genotype likelihoods (settings: --doGlf 2 --doMajorMinor 1 --SNP_pval 1e-6 --doMaf 1). All
221 reference and sample honey bee genomes were analyzed together (total genomes = 130) using
222 14,705,135 variant sites. Genotype likelihoods have been shown to be robust to low-coverage
223 sequencing data (Skotte, Korneliussen, & Albrechtsen, 2013; Kornliussen *et al.,* 2014) such as
224 those of the Wallberg *et al.,* (2014) reference genomes.
225
226 **Admixture and Principal Components Analysis (PCA)**
227 For admixture analysis we used the program NGSadmix (Skotte *et al*., 2013), which uses a
228 genotype-likelihood based approach that factors in uncertainty associated with next-generation

sequencing and has been shown to have good performance even with low-coverage data. We ran NGSadmix using the BEAGLE genotype likelihood files created by ANGSD with K values ranging from 2 to 6 (K = number of assumed genetic clusters). Here we focus on the results from K = 4 genetic clusters because we are interested in assessing the contributions of the four ancestral lineages (A, M, C, and O) historically imported into the Americas. We used R (R Core Team 2014) to graph admixture estimates. We used PCAngsd (Kornliussen *et al.,* 2014) to conduct a principal components analysis of all SNPs, and graphed the resulting PCA using the eigen function in R (R Core Team, 2014).

**Mitochondrial Sequence Assembly and Phylogenetic Analysis**

Filtered reads (described previously) of all 60 sampled honey bees were aligned to a mitochondrial reference genome from an individual of subspecies *Apis mellifera ligustica* sequenced by Crozier & Crozier (1993). We then called variants using samtools v1.10v (mpileup function) and used bcftools v1.10.2 (Li & Durbin, 2009; Li 2011) to extract the consensus sequence and convert to FASTQ with the vcfutils.pl script. We downloaded 12 previously assembled mitochondrial sequences from *A. mellifera* subspecies representing all four major lineages from NCBI to compare with our samples (listed in Table 2).

FASTQ files of mitochondrial sequences from all 73 honey bees (13 reference honey bees and 60 AHB samples) were aligned using MAFFT (Katoh, Rozewicki, & Yamada, 2019), on the XSEDE via Cipres 2.0 Science Gateway. We used MEGAX (Kumar, Stecher, Li, Knyaz, & Tamura, 2018) and complete deletion of gaps and missing data to create a neighbor-joining

phylogeny under a Kimura 2-parameter model to compute evolutionary distances. We then ran 2000 bootstrap replicates to estimate confidence in the resulting phylogeny.

**Measures of Genetic Diversity**

To assess allelic diversity, we calculated estimations of both pairwise theta ($\hat{\theta}_\pi$), based on the number of mean pairwise differences between sequences, and Watterson's theta ($\hat{\theta}_w$), based on the measure of segregating sites for each sampled and reference population using ANGSD v.928 (Kornliussen *et al.,* 2014). Using only sites in which at least 50% of individuals in a population provided data, we estimated the folded site frequency spectrum (SFS) across the entire genome using the reference honey bee genome as the ancestral state. We then calculated and averaged thetas per site, including invariant sites, using ANGSD's realSFS program. To ensure that our diversity estimates were not overly affected by the difference in coverage between our reference and newly-sequenced genomes, we calculated an additional measure of pairwise nucleotide diversity ($\hat{\theta}_\pi$) using only higher-confidence SNPs with >5% minor allele frequency (MAF) in the total sample, following a pipeline described in Calfee *et al.,* (2020). Using ANGSD, we first identified a set of SNPs with > 5% minor allele frequency in the total sample and inferred the major and minor alleles at those SNPs using observed base counts (-doMajorMinor 2 -doCounts 1 -doMaf 8 –minMAF 0.05). We excluded SNPs where more than half of individuals in the total sample did not have coverage. Using this list of SNPs (n = 5,588,252) as a reference, we calculated allele frequencies for each population based on observed base counts in ANGSD (-doMajorMinor 3 -doCounts 1 -doMaf 8). From these population allele frequencies, we calculated the average pairwise diversity per SNP, correcting for small sample sizes. To account for

invariant sites in our estimate of nucleotide diversity ($\pi$) we weighted our measure of $\pi$ per-SNP by the genome SNP density (total number of SNPs / total positions in the genome). For each measure of genome-wide nucleotide diversity, we estimated standard errors by using a block-jackknife procedure, treating each chromosome as a block and re-computing nucleotide diversity with sequential exclusion of each chromosome.

## RESULTS

### Global genomic ancestry in Africanized honey bee samples

As expected, K = 4 clustering clearly separates the 70 reference honey bees from Wallberg *et al.,* (2014) into four major honey bee lineages (A, C, M & O) with limited evidence of admixture between these groups (Figure 1). *Apis mellifera syriaca* (O) is an exception, with ~20% of its ancestry attributed to the African clade, consistent with results found by Wallberg *et al.,* (2014). For additional analyses (K = 2 - 6), see Figure S1. Two individuals from the western European clade (M) (one from subspecies *A. m. mellifera* and one from subspecies *A. m. iberiensis*) showed significant ancestry from other clades (Clades C and O, respectively), a finding also consistent with Wallberg *et al.,* (2014) (Figure 1). We also observed a small proportion of O ancestry (~10%) across all *A. m. scutellata*.

The nuclear genomes of honey bees from Central America and Mexico were heavily Africanized. Honey bees from Panamá averaged 94% (SE 0.23%) African (A) ancestry with the remaining 6%

295 (SE 0.20%) of their genomes derived from the western European (M) lineage. In Costa Rica,
296 honey bees averaged 90% (SE 1.1%) African (A), 6% (SE 0.059%) western European (M) and
297 4% (SE 0.57%) eastern European (C). In Mexico, honey bees averaged 79% (SE 0.62%) African
298 (A), 12% (SE 0.41%) western European (M) and 8% (SE 0.35%) eastern European (C) (Figure 1,
299 Table 2).
300
301 In contrast to the honey bees of Central America and Mexico, genomes of all 15 honeybees
302 sampled from San Diego (California, U.S.A.) exhibited a diverse admixture of all four major
303 clades (A, M, C, and O). Ancestry of San Diego bees averaged 40% (SE 1.2%) African (A), 16%
304 (SE 0.43%) western European (M), 35% (SE 1.2%) eastern European (C) and 9% (SE 0.22%)
305 Middle Eastern (O) (Figure 1). African (A) ancestry of San Diego bees averaged was far lower
306 than that found in bees from any of the other sampled sites and contributions from the eastern
307 European (C) lineage were higher than all other populations sampled. All San Diego bees
308 possessed substantial Middle Eastern (O) ancestry while all other sites sampled had negligible or
309 no ancestry from this clade (Figure 1, Table 2).
310
311 **Principial Component Analysis (PCA)**
312 The principal components analysis of the 70 reference honey bees representing the four major
313 honey bee clades (A, M, C, O) and the 60 honey bees we sampled from Panamá to San Diego
314 separated populations by clade and sampling site (Figure 2). The ancestral honey bee lineages
315 were widely separated from each other on the first two principal component axes. Bees from the
316 four sampled sites (Panamá; Costa Rica; Mexico; San Diego, CA, U.S.A.) separated into distinct
317 clusters with the exception of partial overlap among the bees from Panamá and Costa Rica. Bees

from Mexico, Costa Rica, and Panamá clustered near African (A clade) honey bees. San Diego bees formed a more distant cluster relative to bees from Mexico, Costa Rica, and Panamá, falling more equidistantly between the A, M, C, and O groups, consistent with their ancestry drawing more evenly from all four groups.

**Mitochondrial Ancestry in Africanized honey bee samples**

Each mitochondrial sequence from our sampled honey bees groups strongly with reference mitochondria from one of the four ancestral lineages (A, M, C, O) in a midpoint rooted phylogeny (Figure 3, Table 3). Notably, mitochondrial sequences from subspecies *A. m. anatoliaca* (Clade O) grouped loosely with subspecies *A. m. ligustica* and *A. m. carnica* (both C). *A. m. anatoliaca* has previously been shown to possess C type mitochondria although it remains characterized as an O clade honey bee due to similarities of morphological characters and nuclear markers (Smith, Slaymaker, Palmer, & Kaftanoglu, 1997; Palmer, Smith, & Kaftanoglu, 2000; Wallberg *et al.,* 2014).

**Genetic Diversity**

All four sampled populations have similar levels of genetic diversity and values for admixed AHB populations are consistently higher than those estimated in reference populations (Table 5). Among ancestral lineages, the African lineage is the most diverse, followed by the Middle Eastern (O) lineage, the Western European (M) lineage and lastly, the Eastern European lineage (C).

## DISCUSSION

Africanized honey bee populations exhibit distinct genomic admixture profiles across their Central and North American range, with African ancestry decreasing with increasing latitude (Figures 1 & 2; Table 2). Despite considerable differences among populations, within each population there is little variation in ancestry among individuals. Thus, AHB populations within countries appear to be well-mixed hybrid swarms. Honey bees from Panamá ($\bar{x}$ = 94 % ± 0.23%) and Costa Rica ($\bar{x}$ = 90% ± 1.1%) were the most similar in terms of African ancestry, differing primarily by the presence of small amounts of eastern European (C) ancestry in the Costa Rica sample. While lower than that found in our Costa Rica and Panamá samples, African ancestry in our Mexico sample is also substantial ($\bar{x}$ = 79% ± 0.62%), but unlike the Central American samples, these honey bees possess increased levels of C ($\bar{x}$ = 8% ± 0.35%) and M ancestry ($\bar{x}$ = 12% ± 0.41). Middle Eastern ancestry (O) accounts for only 1% of the ancestry in Mexico ($\bar{x}$ = 1% ± 0.36%).

The substantial amount of C ancestry persisting in honey bees in Mexico suggests that insufficient time may have passed since the arrival of Africanized honey bees (AHB) for honey bees to reach the high African ancestry levels seen in lower latitudes. However, AHB first arrived in southern Mexico in the late 1980s, and studies have shown that levels of African ancestry can reach high, apparently stable, levels in less than a decade (Pinto *et al.*, 2005). Alternatively, the substantial EHB population that existed throughout Mexico prior to AHB arrival could have provided a genetic buffer and allowed for the persistence of C-type despite ample time since

362 contact with AHB (Clarke, Rinderer, Franck, Quezada-Euán, & Oldroyd, 2002). Beekeeping with
363 C-lineage honey bees was widespread across Mexico prior to the arrival of AHB, with an
364 estimated 1.5 million managed colonies present throughout the country (Winston *et al*., 1979; Gu
365 *et al.,* 2002). In contrast, Costa Rica and Panamá both had modest managed beekeeping activity
366 prior to AHB arrival and feral EHB colonies were quite rare, particularly in the rainy lowlands
367 (Roubik & Boreham, 1990; Lobo, 1995). Additionally, many beekeepers in Central America
368 abandoned the trade after AHB arrival and the importation and maintenance of European honey
369 bees diminished substantially (van Veen, Calderon Fallas, Cubero Murillo, & Arce Arce, 1998).
370 Thus, AHB likely encountered a much smaller population of EHB in Central America than in
371 Mexico, allowing for a rapid and extensive Africanization of the honey bee gene pool.
372
373 In striking contrast to the honey bees from Mexico and Central America, African ancestry in
374 honey bees collected in San Diego County, California (U.S.A.) is relatively low ($\bar{x}$ = 40% ±
375 1.2%) with substantial ancestry traceable to all four major honey bee genetic lineages (Figure 1 &
376 2). Surprisingly, all honey bees from the San Diego sample possessed Middle Eastern ancestry
377 (O) ($\bar{x}$ = 9% ± 0.22%). Honey bees from Middle Eastern lineages were only imported to the
378 United States during the last two decades of the 19th century and these limited importations
379 stopped by the beginning of the 20th century (Magnus & Szalanski, 2010). Nevertheless, surveys
380 of ancestry in honey bees in the United States have continued to report the presence of O-clade
381 ancestry in feral honey bees more than a century since their importation ceased (Whitfield *et al*.,
382 2006; Magnus & Szalanski, 2010; Kono & Kohn, 2015, Figures 1 & 2).
383

384 In addition, representation of eastern European (C) ancestry ($\bar{x}$ = 35% ± 1.2%) in San Diego bees
385 is substantially higher than that found in Mexican and Central American samples while the
386 contribution of the M lineage ($\bar{x}$ = 16% ± 0.43%) is also somewhat elevated (Figure 1, Table 2).
387 The large contribution of eastern European (C) ancestry is perhaps indicative of the fact that
388 honey bees from this clade are preferred for agricultural use in the United States. In regions
389 inhabited by feral AHB, European purity of managed honey bee colonies is actively maintained
390 via consistent requeening of colonies with queens from desirable lineages (Schiff & Sheppard,
391 1995, 1996). In the United States, AHB are generally considered unmanageable due to
392 undesirable characteristics such as a higher propensity to sting and to abandon their nests
393 (reviewed in Schneider *et al*., 2004). In contrast, in Mexico, Central and South America, AHB
394 have been largely accepted as the new normal for beekeeping and have been integrated into
395 agricultural work (Ratnieks & Visscher, 1996; Guzman-Novoa & Page, 1999).
396
397 Our findings in San Diego, largely agree with recent WGS studies that assessed African (A),
398 western European (M) and eastern European (C) ancestry in feral honey bees in Southern
399 California (Cridland *et al.*, 2017; Calfee *et al*., 2020). However, we provide the first assessment
400 of Middle Eastern (O) ancestry in Southern California honey bees using whole genome
401 sequencing (WGS). If the Southern California bees sampled in these previous studies contained
402 unassessed genomic content from the O lineage, it was likely assigned as C-type in their analysis,
403 as these clades are the most genetically similar.
404
405 Of particular interest in all of our samples is the persistence of substantial western European (M)
406 ancestry despite African dominance. Studies that have tracked the process of Africanization

elsewhere have shown that African genetic material largely or completely replaces genomic content from the eastern European (C) lineage, while the contribution from the M lineage to genomes of AHB remains substantial and is never completely eliminated (Clarke *et al*., 2002; Pinto *et al*., 2005; Whitfield *et al*., 2006; Cridland *et al*., 2017; Nelson *et al*., 2017). All of our sampled honey bee genomes from San Diego to Panamá possess moderate levels of M ancestry while C ancestry content declines precipitously from north to south and is nearly totally absent in samples from Costa Rica and Panamá. This pattern suggests that the M-lineage content that persists in highly Africanized populations may be selected for while C-lineage content is selected against except where A-lineage contribution declines at higher latitudes (Whitfield *et al*., 2006). Alternatively, small amounts of M ancestry may be neutrally hitchhiking within predominantly African genomes.

Previous studies have identified some regions of M ancestry that appear to be under selection, in particular a region on Chromosome 13 which is associated with a QTL for worker ovary size (Calfee *et al*., 2020; Nelson *et al*., 2017). In addition, some regions associated with nest defense behavior were found to be of western European origin, suggesting that M ancestry is contributing in some way to AHB nest defense, a behavior which has been historically associated with African ancestry (Harpur, Kadri, Orsi, Whitfield, & Zayed, 2020). Future work is needed to determine whether these regions of M ancestry are under selection in our sampled populations.

Despite differences in sampling methods, bees within each sample population were remarkably homogeneous with respect to inferred ancestry. In both Costa Rica and Mexico, bees were sampled from a single site, either a small area of a reserve (Costa Rica) or a single apiary

(Mexico). This limited geographic breadth of sampling could have contributed to the homogeneity of ancestry observed in these sites. However, in Panamá and San Diego, bees were sampled across many tens of kilometers, representing many different colonies, and exhibited similar homogeneity in ancestry.

Mitochondrial analysis of these 4 New World populations is largely consistent with findings from nuclear genomes (Figure 3; Table 4). All bees sampled from Panamá and Costa Rica, where nuclear genomes were predominantly African, carried mitochondria of African origin. In Mexico, the majority of honey bees carried the African mitotype while a few carried the C-type mitochondria. San Diego honey bees harbored a more diverse selection of mitochondrial lineages (A, M and O) with only the C-lineage mitochondria absent in our current sample. However, a previous study of mitochondrial diversity in San Diego County honey bees (Kono & Kohn, 2015) used a larger sample and found mitotypes representing all four clades, with the African mitotype the most frequent (65%) and mitochondria from the other three lineages present in similar proportions. Failure to uncover any mitochondria from the C lineage in the present study likely results from the small number of non-A mitochondria analyzed.

Estimates of genetic diversity among sampled populations are quite similar to one another, but higher than those from Old World reference populations, as has been previously reported for AHBs from other portions of their range (Harpur, Minaei, Kent, & Zayed, *et al.,* 2012; Calfee *et al.,* 2020; Espregueira-Themudo *et al.*, 2020). The African lineage has previously been shown to be more diverse than the European M and C lineages (Harpur *et al*., 2012; Calfee *et al.,* 2020), and we find that the O lineage harbors levels of genetic diversity intermediate between African

453 (A) and European (M and C) lineages (Table 3). Surprisingly, given that the majority of their
454 genomes do not derive from the African lineage, estimates of genetic diversity in our San Diego
455 sample are not markedly lower than those of Mexican and Central American samples. Previous
456 work has shown that admixture between managed European honey bee subspecies increases total
457 genetic diversity beyond that seen in contributing lineages (Harpur *et al.,* 2012). Perhaps the
458 diverse mixture of all four lineages found in San Diego increases the level of genetic diversity—
459 offsetting the effects of reduced African ancestry. Because of their diverse ancestry, the feral
460 honey bee population of San Diego likely contains genes adapted to many different
461 environmental challenges. Their genomes could provide a valuable resource for future attempts
462 to breed desirable traits into managed honey bee populations.
463
464 The complex ancestry of San Diego honey bees results from repeated, human-mediated,
465 introductions of *Apis mellifera* to the New World followed by admixture. This complex
466 admixture may be one factor underlying the tremendous ecological success of feral honey bees in
467 Southern California. Honey bees in San Diego County are responsible for 75% of floral visits to
468 natural vegetation, even at sites within large preserves far from any source of managed honey
469 bees (Hung *et al*., 2018, 2019). This occurs despite the fact that the native bee fauna of Southern
470 California is very diverse with > 600 species recorded in San Diego County alone. While we can
471 find no comparative data on honey bee importance as pollinators in Southern California prior to
472 Africanization, it is at least possible that AHB's diverse genomic ancestry plays a role in their
473 success, as has been suggested for other invasive taxa (Smith *et al.,* 2020).
474

Honey bees sampled in San Diego, CA (and Southern California, more generally) have relatively low African genomic content in comparison to honey bees assessed elsewhere in the United States. In Both Texas and Arizona, levels of African ancestry in feral honey bees are approximately 75% (Whitfield *et al*., 2006; Bozek *et al.,* 2018) similar to levels reported here for Mexico (Figure 2). Potentially, San Diego County possesses a climate more favorable to honey bees of reduced African genomic ancestry in comparison to other regions where African genomic content has been assessed. Models built from climate data at the southern AHB range limit predict that colder winter weather plays a considerable role in halting AHB expansion. (Taylor & Spivak, 1984; Southwick, Roubik, & Williams, 1990; Harrison, Fewell, Anderson, & Loper, 2006). Western San Diego County has a mild Mediterranean climate featuring dry summers with a mean high temperature of 25°C (August). The coldest winter month (January) has an average minimum of 8°C (NOAA - National Weather Service Forecast Office). In contrast, AHB from Texas sampled in Pinto *et al.,* (2005) and reexamined by Whitfield *et al.,* (2006) were collected from the Welder Wildlife Refuge (WWR), a reserve that has a hot and humid summers, on average reaching 35°C in August (Rangel *et al.,* 2016; NOAA - National Weather Service Forecast Office) and experiences cool winters similar to San Diego with average lows of 7°C in January. The greater penetrance of African (A lineage) genes in south Texas, where winter temperatures are, if anything, slightly cooler than San Diego implies that, while climate may be important in limiting the penetrance of African genomic material, simple measures of winter cold temperatures are unlikely to be the only determining factor. Perhaps African ancestry is advantageous in regions that experience much higher summer temperatures and humidity levels. If so, we might expect an increase in Africanization of feral (non-managed) bees with increasing temperatures under climate change.

Gene flow from managed European honey bee populations could restrain the introgression of genes of African origin in San Diego County. Policies throughout the United States are meant to keep managed honey bees as free as possible from Africanization, but such practices have failed to noticeably inhibit the Africanization of feral bees in Texas and Arizona (Pinto *et al.,* 2005; Whitfield *et al.,* 2006, Bozek *et al*., 2018. San Diego County, CA (USA) may differ from other areas in the U.S.A. where AHB ancestry has been assessed in that it has a substantial agricultural component: ~230,000 acres of planted crops, many of which (e.g. avocados and citrus) use honey bees for pollination services and honey production (San Diego County Crop Statistics Annual Report, 2019). San Diego county may harbor higher densities of managed, European honey bees and gene flow from European managed hives could counter Africanization. Genetic swamping by managed honey bees, however, would require that a substantial fraction of the honey bees in San Diego County come from managed, genetically European, hives. Our finding that all 15 foraging workers examined here have substantial African and Middle Eastern ancestry—lineages not used in managed colonies—argues against this. We found no purely European bees in our sample, as would be expected if sampled honey bees came from hives managed in accord with current policies. This is consistent with the hypothesis, supported by previous mitochondrial data (Kono & Kohn, 2015), that most bees foraging in San Diego County, whether in urban or non-agricultural rural settings, derive from feral, Africanized colonies. This high frequency of feral Africanized bees in the total honey bee population reduces the possibility that continued gene flow from managed population is a major force responsible for the low levels of African ancestry observed. The homogeneity of ancestry among sampled bees also suggests that gene flow from

520 managed EHB is relatively rare. We observed little variation indicative of recent introgression
521 events between managed EHB and feral AHB.
522
523 Alternatively, insufficient time may have elapsed since the introduction of the AHB to San Diego
524 county for African ancestry to reach levels comparable to those seen elsewhere. This seems
525 unlikely given the speed with which Africanization has occurred elsewhere. For instance, Pinto *et*
526 *al.*, (2005) used nuclear microsatellite data to show that the transition to high African ancestry in
527 feral bees of southern Texas took only about 5 years after the arrival of AHB into Texas, a
528 finding later confirmed using genomic methods (Whitfield *et al.,* 2006). AHB arrived in San
529 Diego county in 1994 and our bees were sampled more than two decades later, suggesting either
530 that Africanization is taking much longer than in Texas, or differences in conditions in San Diego
531 relative to Texas lead to reduced penetration of genetic material of African origin. It appears that
532 none of these three hypotheses (limited time, gene flow from managed hives, nor low winter
533 temperature) can easily explain lower African content in the genomes of AHB of Southern
534 California. More work is needed to examine whether low African ancestry persists over the long
535 term and what its causes are.
536
537 The introduction of African honey bees to the New World in 1956, and their subsequent
538 introgression and rapid expansion throughout much of the Americas, has captured the attention
539 and imagination of both the scientific world and the public. The Africanization of New World
540 honey bees has been one of the largest and best-documented biological invasions resulting from
541 human-mediated hybridization. The increasing amount of genomic and computational tools
542 available to assess ancestry in hybrid individuals, as well as the ever-decreasing costs of

543 sequencing, have facilitated our ability to study Africanization in unprecedented detail. Here we
544 assessed global ancestry of 60 admixed honey bee genomes collected from four distinct regions,
545 many assessed for the first time using whole genome sequences. Future work to determine local
546 ancestry could investigate whether there are particular genomic regions that consistently come
547 from African versus European lineages across the northern geographic range of the AHB
548 expansion. Such regions, and the genes they contain, are critical to understanding the genetic
549 changes that explain the ecological dominance of Africanized bees over much of the American
550 continents, as well as the continued prevalence of bees of European descent at higher latitudes.
551 Such analyses could also shed light on the locations and origins of genomic regions useful for
552 breeding managed honey bees that are more resistant to environmental challenges currently
553 harming the honey bee industry.
554

555 **ACKNOWLEDGEMENTS**

556

557 We thank Dr. Daniel Sánchez and Dr. Daniel Janzen for collecting honey bees in Chiapas,
558 Mexico, and Guanacaste Costa Rica, respectively. Meg Duell, David Roubik, and William
559 Wcislo provided support and guidance to D.Z during sampling in Panamá. James Nieh and David
560 Holway provided helpful support and feedback during all facets of this research. This research
561 was supported by a FISP grant from UCSD to J.R.K and R.S.B as well as a UC MRPI grant to
562 J.R.K. T.G.L. was supported by NSF PRFB Award no. 1523543.
563

## CITATIONS


Abbott, R., Albach, D. Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *J. Evol. Biol. 26*, 229-246. doi: 10.1111/j.1420-9101.2012.02599.x

Anderson, E., & Stebbins, G. L. J. (1954). Hybridization as an evolutionary stimulus. *Evolution*. *8*, 378-388.

Barton, N. H. (2001). The role of hybridization in evolution. *Molecular Ecology, 10*, 551-568.

Bilo, B. M., Rueff, F., Mosbech, H., Bonifazi, F., Oude-Elberink, J. N. G., & the EACCI Interest Group on Insect Venom Hypersensitivity. (2005). Diagnosis of Hymenoptera venom allergy. *Allergy*, *60*, 1339-1349.

Bozek, K., Rangel, J., Arora, J., Tin, M. M. Y., Crotteau, E., Loper, G. M., … Mikheyev, A. S. (2018). Parallel genomic evolution of parasite tolerance in wild honey bee populations. bioRxiv preprint doi: https://doi.org/10.1101/498436

Breed, M. D., Guzmán-Novoa, E., & Hunt, G. J. (2004). Defensive Behavior of Honey Bees: Organization, Genetics, and Comparisons with Other Bees. *Annual Review of Entomology, 49*, 271-278. doi: 10.1146/annurev.ento.49.061802.123155

Burton, R. S. & Barreto, F. S. (2012). A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities? *Molecular Ecology, 21*, 4942-4957.

Calfee, E., Agra, M. N., Palacio, M. A., Ramirez, S. R., & Coop, G. (2020). Selection and hybridization shaped the Africanized honey bee invasion of the Americas. bioRxiv: 2020.2003.2017.994632. doi: 10.1101/2020.03.17.994632

Clarke, K. E., Rinderer, T. E., Franck, P., Quezada-Euán, J. G., & Oldroyd, B. P. (2002) The Africanization of honey bees (Apis mellifera L.) of the Yucatan: a study of a massive hybridization event across time. *Evolution*, *56,* 1462-1474.

Collins, A. M., Rinderer, T. E., Harbo, J. R., & Bolten, A. B. (1982). Colony defense by Africanized and European honey bees. *Science, 218*, 72-74.

Cridland, J. M., Ramirez, S. R., Dean, C. D., Sciligo, A., & Tsutsui, N. D. (2017). Genome sequencing of museum specimens reveals rapid changes in the genetic composition of honey bees in California. *Genome Biology and Evolution, 10*(2), 458-472.

606 Crozier, R. H., & Crozier, Y. C. (1993). The Mitochondrial Genome of the Honeybee *Apis*
607 *mellifera*: Complete Sequence and Genome Organization. *Genetics*, *13,* 97-117.
608
609 Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2015). jModelTest 2: more models, new
610 heuristics and high-performance computing. *Nature Methods*, *9*(8), 1-4.
611
612 DeGrandi-Hoffman, G., Collins, A., Martin, J. H., Schmidt, J. O., & Spangler, H. G. (1997). Nest
613 defense behavior in colonies from crosses between Africanized and European honey bees (*Apis*
614 *mellifera* L.) (Hymnoptera: Apidae). *Journal of Insect Behavior, 11*(1), 37-45.
615
616 Del Lama, M. A., Lobo, J. A., Soares, A. E. E., & Del Lama, S. N. (1990). Genetic differentiation
617 estimated by isozymic analysis of Africanized honeybee populations from Brazil and from
618 Central America. *Apidologie*, *21*, 271-280.
619
620 Department of Agriculture Weights and Measures. County of San Diego. 2019 Crop Statistics
621 and Annual Report. (2019).
622
623 Dittrich-Reed, D. R. & Fitzpatrick, B. M. (2013). Transgressive hybrids as hopeful monsters.
624 *Journal of Evolutionary Biol*ogy, *40*, 310-315.
625
626 Dobzhansky, T. (1936). Studies on hybrid sterility. II. Localization of sterility factors in
627 *Drosophila psuedoobscura* hybrids. *Genetics*, *21***,** 113-135.
628
629 Fewell, J. H., & Bertram, S. M. (2002). Evidence for genetic variation in worker task
630 performance by African and European honey bees. *Journal of Behavioral Ecology and*
631 *Sociobiology, 52*(4), 318-325.
632
633 Grabenstein, K. C. & Taylor, S. A. (2018). Breaking Barriers: Causes, Consequences, and
634 Experimental Utility of Human-Mediated Hybridization. *Trends in Ecology & Evolution*, *33*,
635 198-212.
636
637 Grant, V. (1981). *Plant speciation.* New York City, NY: Columbia University Press.
638
639 Gu, G., Zhang, C., & Hu, F. (2020, December 11). Analysis on the structure of honey production
640 and trade in the world. [Report]. Retrieved from:
641 http://www.apimondiafoundation.org/foundation/files/
642 2002/GU%20G.%20ZHANG%20CH.pdf
643
644 Guzman-Novoa, E., Sanchez, A., Page Jr., R. E., Garcia, T. (1996). Susceptibility of European
645 and Africanized honeybees (*Apis mellifera* L) and their hybrids to *Varroa jacobsoni* Oud.
646 *Apidologie, 27***,** 93-103.
647
648 Guzman-Novoa, E., & Page Jr., R. E. (1999). Selective breeding of honey bees (Hymenoptera:
649 Apidae) in Africanized areas. *Journal of Economic Entomology, 92(*3), 521-525.
650

Espregueira Themudo, G., Rey-Iglesia, A., Robles Tascón, L., Bruun Jensen, A., da Fonseca, R. R., & Campos, P. F. (2020). Declining genetic diversity of European honeybees along the twentieth century. *Science Rep*orts, **1***0*, 10520. https://doi.org/10.1038/s41598-020-67370-2

Hall, H. G., & McMichael, M. A. (2001). Frequencies of restriction fragment-length polymorphisms indicate that neotropical honey bee (Hymenoptera: Apidae) populations have African and west European origins. *Annals of the Entomological Society of America, 94*, 670–676.

Harpur, B., Minaei, S., Kent, C. F., & Zayed, A. (2012). Management increases genetic diversity of honey bees via admixture. *Molecular Ecology*, *21,* 4414 – 4421.

Harpur, B., Kadri, S. M., Orsi, R. O., Whitfield, C. W., & Zayed, A. (2020). Defense response in Brazilian honey bees (*Apis mellifera scutellata x spp*.) is underpinned by complex patterns of admixture. *Genome Biology and Evolution*, evaa128, https://doi.org/10.1093/gbe/evaa128

Harrison, J. F., Fewell, J. H., Anderson, K. E. & Loper, G. M. (2006). Environmental physiology of the invasion of the Americas by Africanized honeybees. *Society for Integrative and Comparative Biology, 46*, 1110–1122.

Hedrick, P. (2013). Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, *22*, 4606-4618.

Huerta-Sánchez, E., Jin, X., Bianba, A., Bianba, Z., Peter, B. J., Vinckenbosch, N., … Nielsen, R. (2014). Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature*, *512*, 194-197.

Hung, K-L. J., Kingston, J. M., Albrecht, M., Holway, D., & Kohn, J. R. (2018). The worldwide importance of honey bees as pollinators in natural habitats. *Proc. R. Soc. B, 285:* 20172140. http://dx.doi.org/10.1098/rspb/2017.2140

Hung, K-L. J., Kingston, J. M., Lee, A., Holway, D., & Kohn, J. R (2019). Non-native honey bees disproportionately dominate the most abundant floral resources in a biodiversity hotspot. *Proc. R. Soc. B, 286,* 20182901. http://dx.doi.org/10.1098/rspb/2018.2901

Katoh, K, Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics, 20*(4), 1160- 1166.

Kim KT, Oguro J. Update on the status of Africanized honey bees in the Western states. *West J Med* 1999; 1 70:220-222)

693 Kofler, R., Orozco-terWengel, P., De Maio, N., Pandey, R. V., Nolte, V., Futschik, A., …
694 Schlötterer, C. (2011). PoPoolation: A Toolbox for Population Genetic Analysis of Next
695 Generation Sequencing Data from Pooled Individuals. *PLoS ONE, 6*(1), e15925.
696
697 Kono, Y., & Kohn, J. R. (2015). Range and frequency of Africanized honey bees in California
698 (USA). *Plos One.*
699
700 Korneliussen, T. S., Albrechtsen, A., & Nielsen, R., (2014). ANGSD: Analysis of Next
701 Generation Sequencing Data. *BMC Bioinformatics, 15,*356. http://www.biomedcentral.com/1471-
702 2105/15/356
703
704 Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular
705 Evolutionary Genetics Analysis across computing platforms. *Molecular Biology & Evolution*, *35*,
706 1547-1549.
707
708 Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping
709 and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*(21):
710 2987–2993.
711
712 Li, H., & Durbin, R. (2009a). Fast and accurate long-read alignment with Burrows-Wheeler
713 transform. *Bioinformatics*, *26*(5), 589-595.
714
715 Li, H., Handsaker, B., Wysocker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R., and 1000
716 Genome Project Data Processing Subgroup. (2009). The Sequence alignment/map (SAM) format
717 and SAMtools. *Bioinformatics, 25*, 2078-2079. doi:10.1093/bioinformatics/btp352
718
719 Lin W., McBroome, J., Rehman, M., & Johnson, B. R. (2017). Africanized bees extend their
720 distribution in California. *PLoS ONE, 13*(1). e0190604; https://doi.org/10.1371/journal.
721 pone.0190604
722
723 Lobo, J. A. (1995). Morphometric, isozymic, and mitochondrial variability of Africanized
724 honeybees in Costa Rica. *Heredity*, *75,* 153-141.
725
726 Lobo, J. A., Del Lama, M. A., & Mestriner, M. A. (1989). Population differentiation and racial
727 admixture in the Africanized honey bee (Apis mellifera L.) *Evolution*, 43, 794-802.
728
729 Magnus, R., & Szalanski, A. L. (2010). Genetic Evidence for Honey Bees (*Apis mellifera* L.) of
730 Middle Eastern Lineage in the United States. *Sociobiology, 55*(1), 285-296.
731
732 Mayr, E. (1942). *Systematics and the Origin of Species.* New York City, NY. Columbia
733 University Press.
734
735 McNally, L. C., & Schneider, S. S. (1992a). Factors influencing seasonal absconding in colonies
736 of the African honey bee, *Apis mellifera scutellata. Insectes Sociaux, 39,* 403-423.

McNally, L. C., & Schneider, S. S. (1992b). Seasonal cycles of growth, development and movement of the African honey bee, *Apis mellifera scutellata*, in Africa. *Insectes Sociaux*, *39*, 167–179.

McNally, L. C., & Schneider, S. S. (1996). Spatial distribution and nesting biology of colonies of the African honey bee *Apis mellifera scutellata* (Hymenoptera: Apidae) in Botswana, Africa. *Environmental Entomology, 25*, 643–652.

Morse, R. A., & Calderone, N. W. (2020, December 10). The value of honey bees as pollinators of U.S. crops in 2000. [Report]. Retrieved from: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.554.5898&rep=rep1&type=pdf

Muller, H. J. (1942). Isolating mechanisms, evolution and temperature. In T. Dobzhansky (Ed.), *Biological Symposia: A Series of Volumes Devoted to Current Symposia in the Field of Biology* (Vol. 6). (pp. 71-125). Lancaster, PA: Jacques Cattell Press.

National Weather Service Forecast Office. (2020 November 12). National Oceanic and Atmospheric Administration (NOAA). [Database]. Retrieved from W2.weather.gov.

Nelson, R. M., Wallberg, A., Simoes, Z. L. P., Lawson, D. J., & Webster, M. T. (2017). Genomewide analysis of admixture and adaptation in the Africanized honeybee. *Molecular Ecology*, *26*, 3603-3617.

Otis, G. W. (1991). Population biology of the Africanized honey bee. In: M. Spivak, D. J. C. Fletcher, & M. D. Breed, (Eds.), *The African Honey Bee* (pp. 213–34). Boulder, CO: Westview Press.

Palmer, M. R., Smith, D. R., & Kaftanoglu, O. (2000). Turkish Honeybees: Genetic Variation and Evidence for a Fourth Lineage of Apis Mellifera MtDNA. *Journal of Heredity, 91*, 42 – 66.

Pinto, M. A., Rubink, W. L., Patton, J. C., Coulsen, R. N., & Johnston, J. S. (2005). Africanization in the United States: Replacement of Feral European Honeybees (*Apis mellifera* L.) by an African Hybrid Swarm. *Genetics, 170*, 1653-1665.

Quezada-Euan, J. J. G. & Hinsull, S. M. (1995). Evidence of continued European morphometrics and mtDNA in feral colonies of honey bees (*Apis mellifera*) from the Yucatán peninsula, Mexico. *Journal of Apicultural Research*, *34*(3), 161-166.

QuickStats: Number of Deaths from Hornet, Wasp, and Bee Stings, Among Males and Females — National Vital Statistics System, United States, 2000–2017. MMWR Morb Mortal Wkly Rep 2019;68:649. DOI: http://dx.doi.org/10.15585/mmwr.mm6829a5external icon.

R Core Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rabe, M. J., Rosenstock, S. S., & Nielsen, D. I. (2005). Feral Africanized Honey Bees (Apis mellifera) in Sonoran Desert Habitats of Southwestern Arizona. *The Southwestern Naturalist,. 50(*3), 307-311.

Rangel, J., Giresi, M., Pinto, M. A., Baum, K. A., Rubink, W. L., Coulsen, R. N., & Spencer Johnston, J. (2016). Africanization of a feral honey bee (*Apis mellifera*) population in South Texas: does a decade make a difference? *Ecology and Evolution, 6,* 2158-2169.

Ratnieks, F., & Visscher, P. K. (2020, December 11). Living with the Africanized bee: Sinaloan beekeepers adapt pollination to Africanized bees. [Report]. Retrieved from https://doi.org/10.3733/ca.v050n04p24

Rieseberg, L. H., Raymond, O., Rosenthal, D. M., Lai, Z., Livingstone, K., Nakazato, T., … Lexer, C. (2003). Major ecological transitions in wild sunflowers facilitated by hybridization. *Science, 301*, 1211-1216.

Roubik, D. W., & Boreham, M. M. (1990). Learning to Live with Africanized Honeybees. *Interciencia*, *15*(3), 146-153.

Rubink, W. L., Luévano-Martinez, P., Sugden, E. A., Wilson., W. T., & Collins, A. M. (1996). Subtropical *Apis mellifera* (Hymnoptera: Apidae) swarming dynamics and Africanization rates in northeastern Mexico and southern Texas. *Annals of the Entomological Society of America, 89*, 243-251.

Ruttner, F. (1988). *Biogeography and Taxonomy of Honeybees.* New York City, NY: Springer.

Schneider, S. S., DeGrandi-Hoffman, G., & Smith, D. R. (2004). The African Honey Bee: Factors Contributing to a Successful Biological Invasion. *The Annual Review of Entomology, 49*, 351-376.

Schiff, N. M. & Sheppard, W. S. (1993). Mitochondrial DNA evidence for the 19$^{th}$ century introduction of African honey bees into the United States. *Experientia*, *49,* 350-52.

Schiff, N. M., & Sheppard, W. S. (1995). Genetic analysis of commercial honey bees (Hymenoptera: Apidae) from the southern United States. *Journal of Economic Entomology, 88*, 1216-1220.

Schiff, N. M., & Sheppard, W. S. (1996). Genetic differentiation in the queen breeding population of the western United States. *Apidologie*, *27*, 77-86.

Schiff, N. M., Sheppard, W. S., Loper, G. M., Shimanuki, H. (1994). Genetic Diversity of Feral Honey Bee (Hymeoptera: Apidae) Populations in the Southern United States. *Annuls of the Entomological Society*, *87*(6): 842-848.

Seehausen, O. (2004) Hybridization and adaptive radiation. *Trends in Ecology and Evolution, 19*, 198-207.

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics, 195,* 693-702.

Smith, A. L., Hodkinson, T. R., Villellas, J., Catford, J. A., Csergo, A. M., Blomberg, S. P., … Buckley, Y. M. (2020). Global gene flow releases invasive plants from environmental constraints on genetic diversity. *PNAS, 117*(8): 4218-4227.

Smith, D. R., Taylor, O. R., Brown, W. M. (1989). Neotropical Africanized honey bees have African mitochondrial DNA. *Nature*, *339*, 213-215.

Smith, D. R., Slaymaker, A., Palmer, M., & Kaftanoglu, O. (1997). Turkish Honey Bees Belong to the East Mediterranean Mitochondrial Lineage. *Apidologie*, 2*8*, 269 - 74.

Southwick, E. E., Roubik, D. W., & Williams, J. M. (1990). Comparative energy balance in groups of Africanized and European honey bees: ecological implications. *Comparative Biochemistry and Physiology, 97*, 1–7.

Stamatakis, A. (2014). RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics*, *1;30*(9), 1312-1313. DOI: 10.1093/bioinformatics/btu033

Stebbins, G. L. (1950). *Variation and evolution in plants*. New York City, NY: Columbia University Press.

Suarez-Gonzalez, A., Lexer, C., & Cronk, Q. C. B. (2018). Adaptive introgression: a plant perspective. *Biology Letters*, *14***.** http://dx.doi.org/10.1098/rsbl.2017.0688

Taylor, O. R., & Spivak, M. (1984). Climatic limits of tropical African honey bees in the Americas. *Journal of the Kansas Entomologicla Society, 53*, 157-165.

The HoneyBee Genome Sequencing Consortium. (2006). Insights into social insects from the genome of the honeybee *Apis mellifera. Nature*, *443*(7714), 931-949.

van Veen, J., Calderon Fallas, R. A., Cubero Murillo, A., & Arce Arce, H. G. (1998) *Varroa jacobsoni* in Costa Rica: detection, spread and treatment with formic acid. *Bee World, 79*(1), 5-10.

VonHoldt, B., Fan, Z., Vecchyo, D. O., & Wayne, R. K. (2017). *EPAS1* variants in high altitude Tibetan wolves were selectively introgressed into highland dogs. *PeerJ.* 5:e3522; DOI 10.7717/peerj.3522

Wallberg, A., Han, F., Wellhagen, G., Dahle, B., Kawata, M., Haddad, N., … Webster, M. (2014). A worldwide survey of genome sequence variation provides insight into the evolutionary history of the honeybee *Apis mellifera. Nature Genetics, 46*(10), 1081-1088.

Whitfield, C. W., Behura, S. K., Berlocher, S. H., Clark, A. G., Spencer Johnston, J., … Tsutsui, N. D. (2006). Thrice Out of Africa: Ancient and recent expansions of the honey bee, *Apis mellifera*. *Science, 314*, 642-645.

Whitney, K. D., Broman, K. W., Kane, N. C., Hovic, S. M., Randell, R. A., & Rieseberg, L. H. (2015). Quantitative trait locus mapping identifies candidate alleles involved in adaptive introgression and range expansion in a wild sunflower. *Molecular Ecology*, *24*, 2194-2211.

Winston, M. L. (1992). The Biology and Management of Africanized Honey Bees. *Annual Review of Entomology, 37*, 173-193.

Winston, M. L. (1979). The potential impact of the Africanized honey bee on apiculture in Mexico and Central America. *American Bee Journal, 119*, 584–586.

**Data Accessibility and Benefit-Sharing Statement**

Upon acceptance, all DNA sequence data will be deposited in GenBank and accession numbers provided. All scripts used to generate analyses will be made available on GitHub.

**Author Contributions**

Grants to J.R.K., R.S.B, D.Z., and T.L. supported the research. D.Z and J.R.K designed the project. D.Z. performed sampling, DNA preparation and data analysis. E.C., T. L. and J.P and J.R.K. aided in data analysis. D.Z. and J.R.K. were principal authors of the paper with the help from all other authors.

## FIGURES

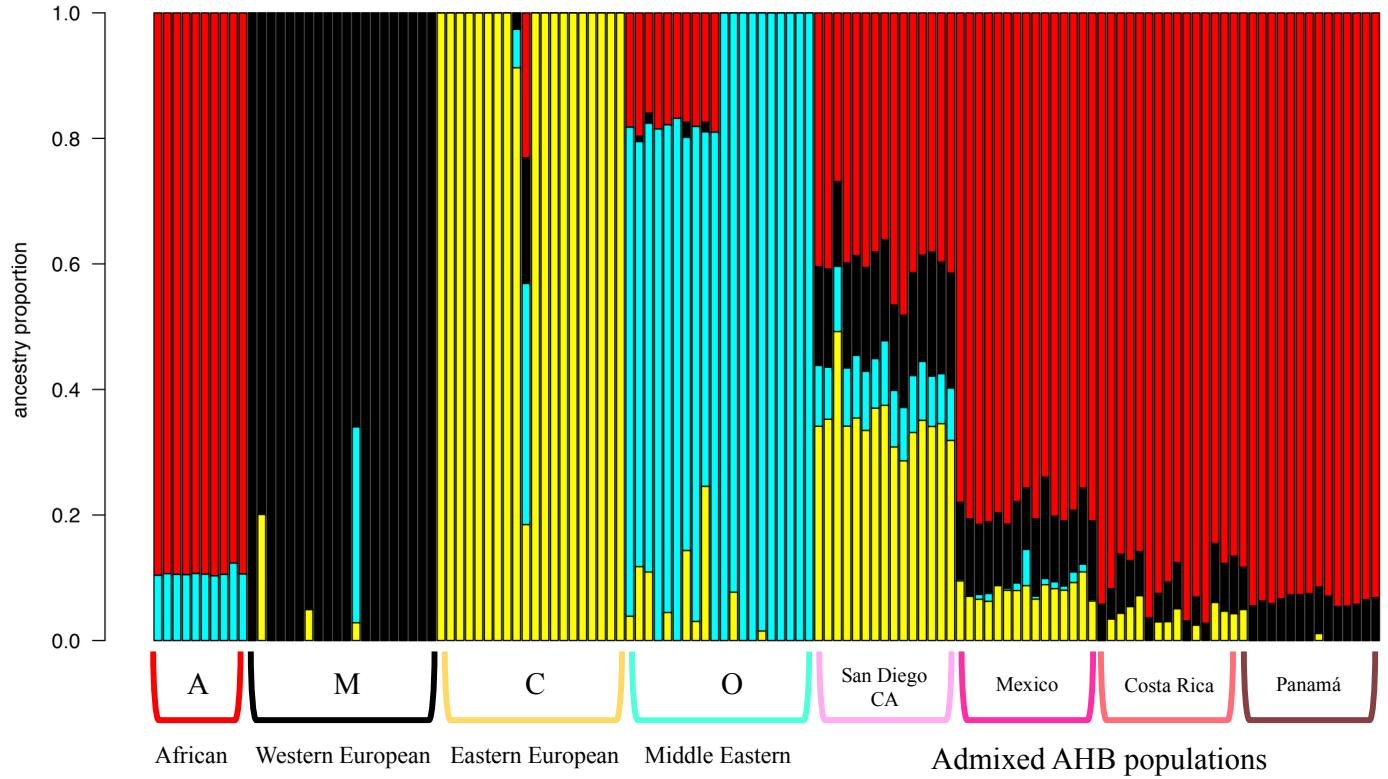


Figure 1: NGSadmix barplot of ancestry. Each vertical bar is one honey bee genome and colors represent the estimated proportion of ancestry derived from each genetic cluster (K=4). The 70 reference genomes belonging to the four major evolutionary lineages of *Apis mellifera* (A, M, C, O) are grouped and labeled beginning with the African clade. The 60 admixed AHB genomes are arranged north to south by geographic origin, beginning with San Diego, CA and followed by the honey bees from Mexico, Costa Rica, and Panamá.

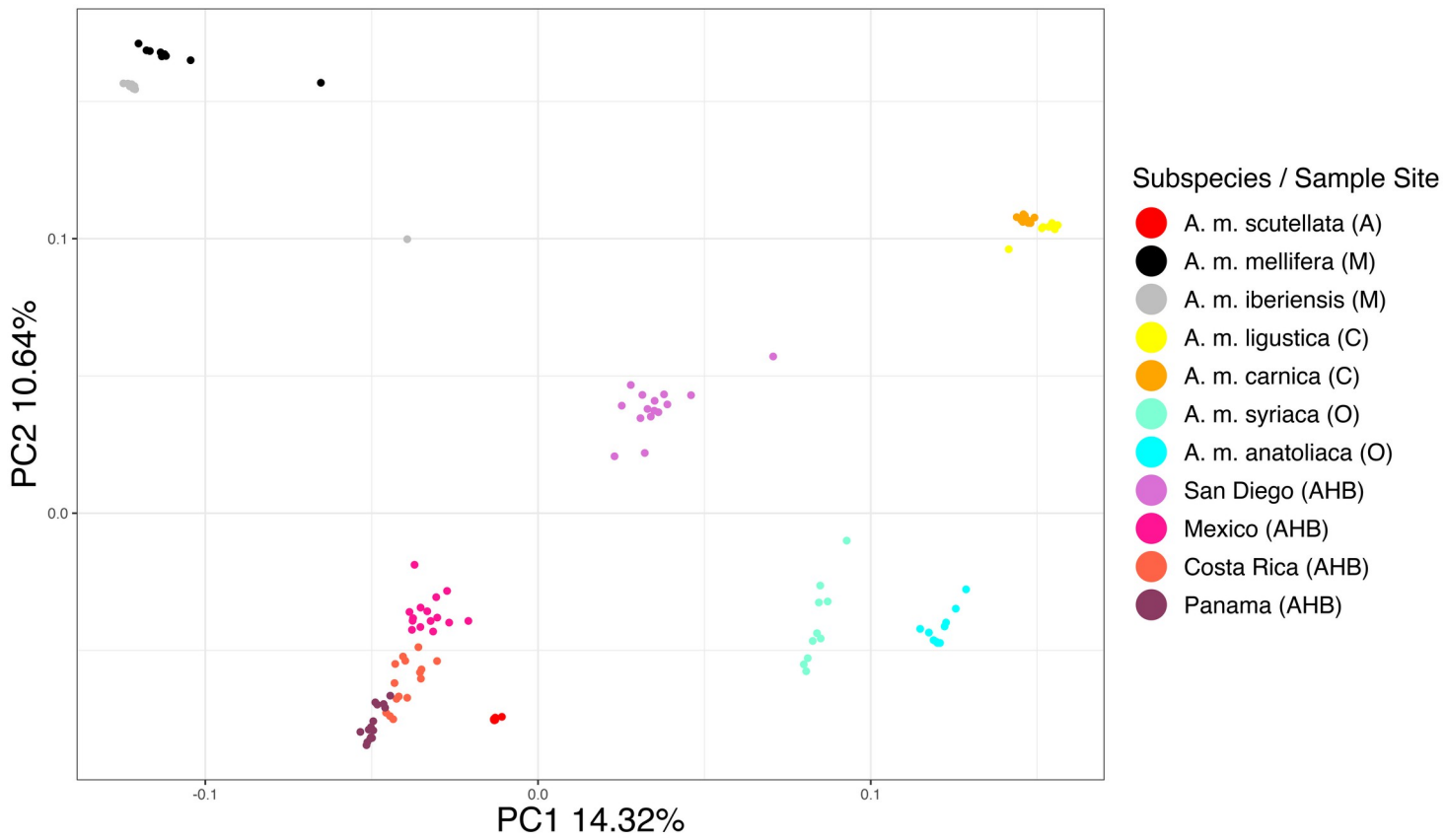


Figure 2: Principal Component Analysis (PCA) of the 70 reference honey bees and 60 admixed honey bee genomes.

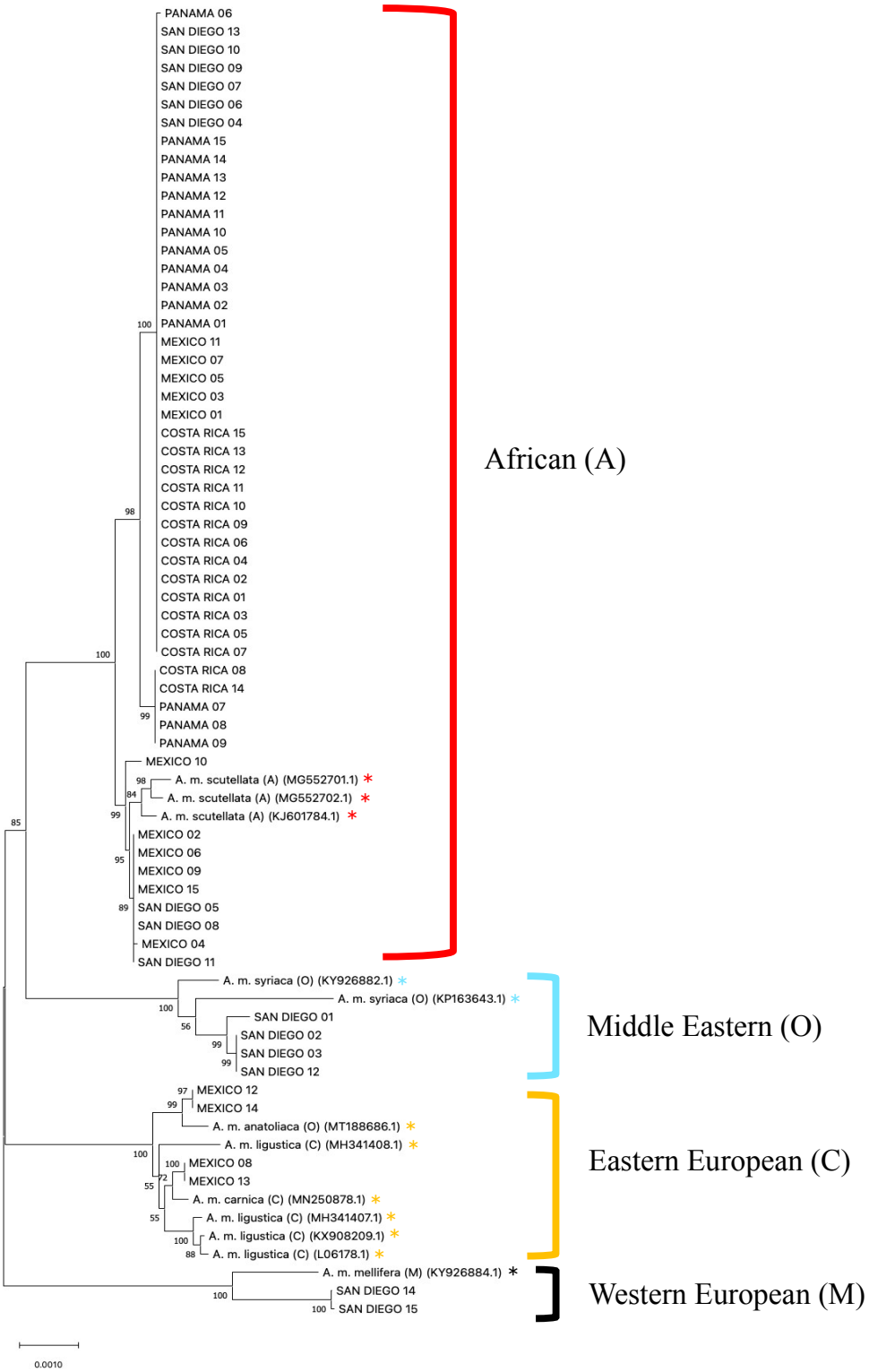
PANAMA 06
SAN DIEGO 13
SAN DIEGO 10
SAN DIEGO 09
SAN DIEGO 07
SAN DIEGO 06
SAN DIEGO 04
PANAMA 15
PANAMA 14
PANAMA 13
PANAMA 12
PANAMA 11
PANAMA 10
PANAMA 05
PANAMA 04
PANAMA 03
PANAMA 02
PANAMA 01
MEXICO 11
MEXICO 07
MEXICO 05
MEXICO 03
MEXICO 01
COSTA RICA 15
COSTA RICA 13
COSTA RICA 12
COSTA RICA 11
COSTA RICA 10
COSTA RICA 09
COSTA RICA 06
COSTA RICA 04
COSTA RICA 02
COSTA RICA 01
COSTA RICA 03
COSTA RICA 05
COSTA RICA 07
COSTA RICA 08
COSTA RICA 14
PANAMA 07
PANAMA 08
PANAMA 09
MEXICO 10
A. m. scutellata (A) (MG552701.1)
A. m. scutellata (A) (MG552702.1)
A. m. scutellata (A) (KJ601784.1)
MEXICO 02
MEXICO 06
MEXICO 09
MEXICO 15
SAN DIEGO 05
SAN DIEGO 08
MEXICO 04
SAN DIEGO 11
A. m. syriaca (O) (KY926882.1)
A. m. syriaca (O) (KP163643.1)
SAN DIEGO 01
SAN DIEGO 02
SAN DIEGO 03
SAN DIEGO 12
MEXICO 12
MEXICO 14
A. m. anatoliaca (O) (MT188686.1)
A. m. ligustica (C) (MH341408.1)
MEXICO 08
MEXICO 13
A. m. carnica (C) (MN250878.1)
A. m. ligustica (C) (MH341407.1)
A. m. ligustica (C) (KX908209.1)
A. m. ligustica (C) (L06178.1)
A. m. mellifera (M) (KY926884.1)
SAN DIEGO 14
SAN DIEGO 15
0.0010
African (A)
Middle Eastern (O)
Eastern European (C)
Western European (M)

Figure 3: Midpoint-rooted neighbor joining phylogeny constructed from the mitochondrial genomes of 60 admixed honey bees collected from San Diego, Mexico, Costa Rica, and Panamá (n=15, per population) and 12 reference mitochondrial sequences obtained from NCBI: *A. m. mellifera* (n=1), *A. m. syriaca* (n=2), *A. m. carnica* (n=1), *A. m. scutellata* (n=3), *A. m. ligustica* (n=4), and *A. m. anatoliaca* (n=1). NCBI mitochondrial sequences denoted by an asterisk (*). Values on each node represent the percent bootstrap support (n = 2000 bootstraps).

## TABLES

**A**

| Clade | Subspecies | (n) | Source Country |
|---|---|---|---|
| A | *A. m. scutellata* | 10 | South Africa |
| M | *A. m. mellifera* | 10 | England |
| | *A. m. iberiensis* | 10 | Ireland |
| C | *A. m. carnica* | 10 | Italy |
| | *A. m. ligustica* | 10 | Greece |
| O | *A. m. syriaca* | 10 | Syria |
| | *A. m. anatoliaca* | 10 | Lebanon |

**B**

| Location | (n) | Coordinates |
|---|---|---|
| San Diego, CA, U.S.A. | 15 | 32.7° N, 117° W |
| Chiapas, Mexico | 15 | 16.7° N, 93.1° W |
| Santa Rosa National Park, Costa Rica | 15 | 10.8° N, 85.7° W |
| Panamá | 15 | 8.98°N, 79.5° W |

Table 1: Summary of all 130 genomes included in this ancestry analysis, including (A) 70 reference honey bee genomes downloaded from NCBI from Wallberg *et al.,* (2014). (B) 60 admixed honey bee genomes collected from four distinct sampling sites.

| | San Diego, CA | Mexico | Costa Rica | Panamá |
|---|---|---|---|---|
| African (A) | 39.7 % ± 1.22 % | 79.2% ± 0.618 % | 90.2 % ± 1.10 % | 93.5 % ± 0.233 % |
| Western European (M) | 16.3 % ± 0.427 % | 11.7%± 0.414 % | 6.23 % ± 0.586 % | 6.46 % ± 0.196 % |
| Eastern European (C) | 35.0 % ± 1.17 % | 8.07% ± 0.347 % | 3.61 % ± 0.566 % | 0.0744 % ± 7.44e-02 % |
| Middle Eastern (O) | 9.03 % ± 0.217 % | 1.08% ± 0.359 % | 1.46e-03 % ± 1.14e-03 % | 0.00 % ± 8.01e-19 % |

Table 2: Mean percentage (SE) of genomic contributions from the four major honey bee lineages in each sampled population (n = 15 bee genomes per sample).

| GenBank Accession Number | Subspecies | Clade |
|---|---|---|
| KJ601784.1 | *A. m. scutellata* | A |
| MG552702.1 | *A. m. scutellata* | A |
| MG552701.1 | *A. m. scutellata* | A |
| KY926884.1 | *A. m. mellifera* | M |
| MN250878.1 | *A. m. carnica* | C |
| KX908209.1 | *A. m. ligustica* | C |
| MH341408.1 | *A. m. ligustica* | C |
| MH341407.1 | *A. m. ligustica* | C |
| L06178.1 | *A. m. ligustica* | C |
| MT188686.1 | *A. m. anatoliaca* | O |
| KP163643.1 | *A. m. syriaca* | O |
| KY926882.1 | *A. m. syriaca* | O |

Table 3: Whole mitochondrial genome sequences representing A/C/M/O honey bee clades downloaded from NCBI and used in mtDNA haplotype analysis.

| | San Diego, CA | Mexico | Costa Rica | Panamá |
|---|---|---|---|---|
| African (A) | 9 | 11 | 15 | 15 |
| Western European (M) | 2 | 0 | 0 | 0 |
| Eastern European (C) | 0 | 4 | 0 | 0 |
| Middle Eastern (O) | 4 | 0 | 0 | 0 |

Table 4: Number of honey bees sampled from each admixed population (San Diego (CA), Mexico, Costa Rica, Panamá) found to carry mitochondria from each of the four clades (A, M, C, O).

| | Pairwise Estimator ($\hat{\theta}_\pi$) | Pairwise Estimator ($\hat{\theta}_\pi$) using called SNPs only (minor allele frequency > 0.05) | Watterson's Estimator ($\hat{\theta}_W$) |
|---|---|---|---|
| San Diego | 0.0101 ±0.00139 | 0.00724 ± 5.40e-05 | 0.0115 ±0.00145 |
| Mexico | 0.0115 ±0.00241 | 0.00767 ± 7.04e-05 | 0.0132 ±0.00178 |
| Costa Rica | 0.0111 ±0.00180 | 0.00745 ± 7.43e-05 | 0.0117 ±0.00194 |
| Panamá | 0.0109 ±0.00186 | 0.00737 ± 7.21e-05 | 0.0121 ±0.00149 |
| African (A) | 0.00842 ±0.000543 | 0.00468 ± 3.10e-05 | 0.0111 ±0.00252 |
| Western European (M) | 0.00438 ±0.000258 | 0.00308 ± 1.33e-05 | 0.00469 ±0.000387 |
| Eastern European (C) | 0.00346 ±0.000195 | 0.00269 ± 1.48e-05 | 0.00369 ±0.000176 |
| Middle Eastern (O) | 0.00642 ±0.000345 | 0.00419 ± 1.91e-05 | 0.00714 ±0.000266 |

Table 5: Genetic diversity measures (mean ± SE) for admixed and reference populations.