Genome Wide Analysis of Admixture in *Apis mellifera intermissa*

Caroline Ritchie

A thesis submitted to the Faculty of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Master of Science

Graduate Program in Biology

York University,

Toronto, Ontario

April 2025

Abstract

Unraveling the evolutionary history of *Apis mellifera* reveals how populations have adapted to diverse environments, resisted disease, and responded to human influence. An admixed population of *A. m. intermissa* in Morocco has sparked debate over the dispersal routes of *A. mellifera* and the origin of the European (M) lineage. Yet, despite its identification, the genomic consequences of this admixture remain largely unexplored. We analyzed global and local ancestry, estimated admixture timing, and assessed genetic diversity in this *A. m. intermissa* population. Our findings reveal recent admixture, occurring ~14 generations ago, with genome-wide diversity reflecting an intermediate value between progenitor lineages. Notable regions on chromosomes 7 and 11 showed high enrichment for M lineage ancestry. Functional enrichment and prior studies suggest these regions influence detoxification, immunity, development, and hormonal regulation. Four QTLs with 18 loci further support M lineage contributions to *Varroa destructor* resilience. These findings reveal the functional role of M lineage ancestry in North African *Apis mellifera* populations and offer a framework for further exploring the genomic outcomes of admixture.

Acknowledgments

I would like to express my deepest gratitude to Dr. Amro Zayed for providing me with the opportunity of a lifetime. Thank you, Dr. Zayed, for being an outstanding mentor. A heartfelt thank you to Dr. Kathleen Dogantzis, whose guidance ignited my passion for researching this particular population. I am immensely grateful to Dr. Lawrence Packer for his willingness to share his wisdom and guidance, and for embodying the true spirit of a dedicated scientist. His mentorship has been a beacon of knowledge and inspiration. Special thanks to Ashwarya, Aiden, Tanushri, Dova, and all the members of the Amro Zayed Lab. Your advice and camaraderie have enriched my experience and have been pivotal in my development as a researcher. I would also like to extend my profound appreciation to my family, whose endless support and encouragement have been my stronghold throughout this journey. And finally, a salute to the incredible and diverse honey bees across the globe, whose fascinating evolutionary tales continue to inspire and intrigue scientists and enthusiasts alike.

Table of Contents

List of Tables

List of Figures

Introduction

The western honey bee, *Apis mellifera*, holds substantial economic and cultural importance, a testament to its enduring relationship with humanity that spans nearly 11,000 years (Batra, 1995; Bloch et al., 2010; Carpenter & Harpur, 2021; Crane, 1999; Roffet-Salque et al., 2015). This long-standing partnership began with the practices of wax and honey collection, expanding more recently to include the role of pollination services (Batra, 1995; Bloch et al., 2010; Carpenter & Harpur, 2021; Crane, 1999; Roffet-Salque et al., 2015). Today, *Apis mellifera* is integral to agricultural industries worldwide, with global honey production estimated to be worth 8.17 billion USD (Fortune Business Insights, 2022), in addition to contributing between 235 billion to 577 billion in annual global food production from pollination services (Bayer Global, 2023), illustrating the continued significance of honey bees through the ages.

Apis mellifera has a large native range across Europe, Africa, and Western Asia, fostering a wide variety of behavioral and morphological variation stemming from historical patterns of isolation, local adaptation, and selective breeding through apicultural practices. *Apis mellifera* has been categorized into seven evolutionary lineages: the M lineage in Eurasia, the C lineage in Eastern Europe, the O and Y lineages in Western Asia, the L lineage in Egypt, the U lineage in Madagascar and the A lineage in Africa. Furthermore, these lineages can be further divided into 29 subspecies (Figure 1) (Cridland et al., 2017; Dogantzis et al., 2021; Han et al., 2012a; Miguel et al., 2011; Ruttner & Ruttner, 1988; Ruttner et al., 1978; Whitfield et al., 2006). These lineages were first described and characterized morphologically by Ruttner (1988) and later validated in genetic studies, initially using mitochondrial DNA and subsequently with whole-genome data (Carpenter & Harpur, 2021; Dogantzis et al., 2021; Toth & Zayed, 2021).

Human management of *Apis mellifera* has facilitated its global distribution, with many instances of honey bee importation documented in efforts to introduce desirable apicultural traits into local populations, resulting in numerous admixed populations. The admixture between *Apis mellifera* lineages, particularly between the European (M lineage) and African (A lineage) bees, has generated considerable research interest (Carpenter & Harpur, 2021; Harpur et al., 2012). These hybrid bees, known as "Africanized honey bees" (AHBs), gained attention in the 1960s when an experimental population of *A. m. scutellata* (A lineage), from South Africa was accidentally released into the Brazilian honey bee population, comprised mainly of M lineage bees, specifically *A. m. iberiensis* (Kent, 1988; Nelson et al., 2017; Scott Schneider et al., 2004). Despite being a genetic blend of European and African populations, the bees predominantly displayed phenotypic traits characteristic of African (A lineage) honey bees, including their nesting biology, swarming behavior, and foraging traits, which contributed to their highly successful biological invasion throughout South America (Nelson et al., 2017). Consequently, AHBs made their way into the United States in the 1990s and have since been documented in 10 southern US states (Carpenter & Harpur, 2021).

Numerous studies focus on understanding the genomic outcomes of AHB hybrids, particularly in North and South America, where their introduction resulted in undesirable behaviors in managed honey bee populations (Calfee et al., 2020; Everitt et al., 2023; Nelson et al., 2017; Wallberg et al., 2014; Winston, 1992). Outside of the Americas, the African hybridization within Spain, specifically around the Iberian Peninsula, also became a subject of great interest when it was discovered that *A. m. iberiensis* (M lineage) exhibited features of A

lineage hybridization detected through mitochondrial DNA (Cánovas et al., 2008; Chávez-Galarza et al., 2015; Hepburn & Radloff, 2013).

This thesis work focuses on a unique African-European hybrid population of *A. m. intermissa* sampled throughout Morocco (Figure 1). These samples of *A. m. intermissa* were originally collected for phylogenetic analysis, and were subsequently identified as an admixed population (Dogantzis et al., 2021; Whitfield et al., 2006). As a result, this *A. m. intermissa* population presents an ideal model for advanced admixture analysis, providing an opportunity to investigate the genomic consequences of admixture in North Africa.

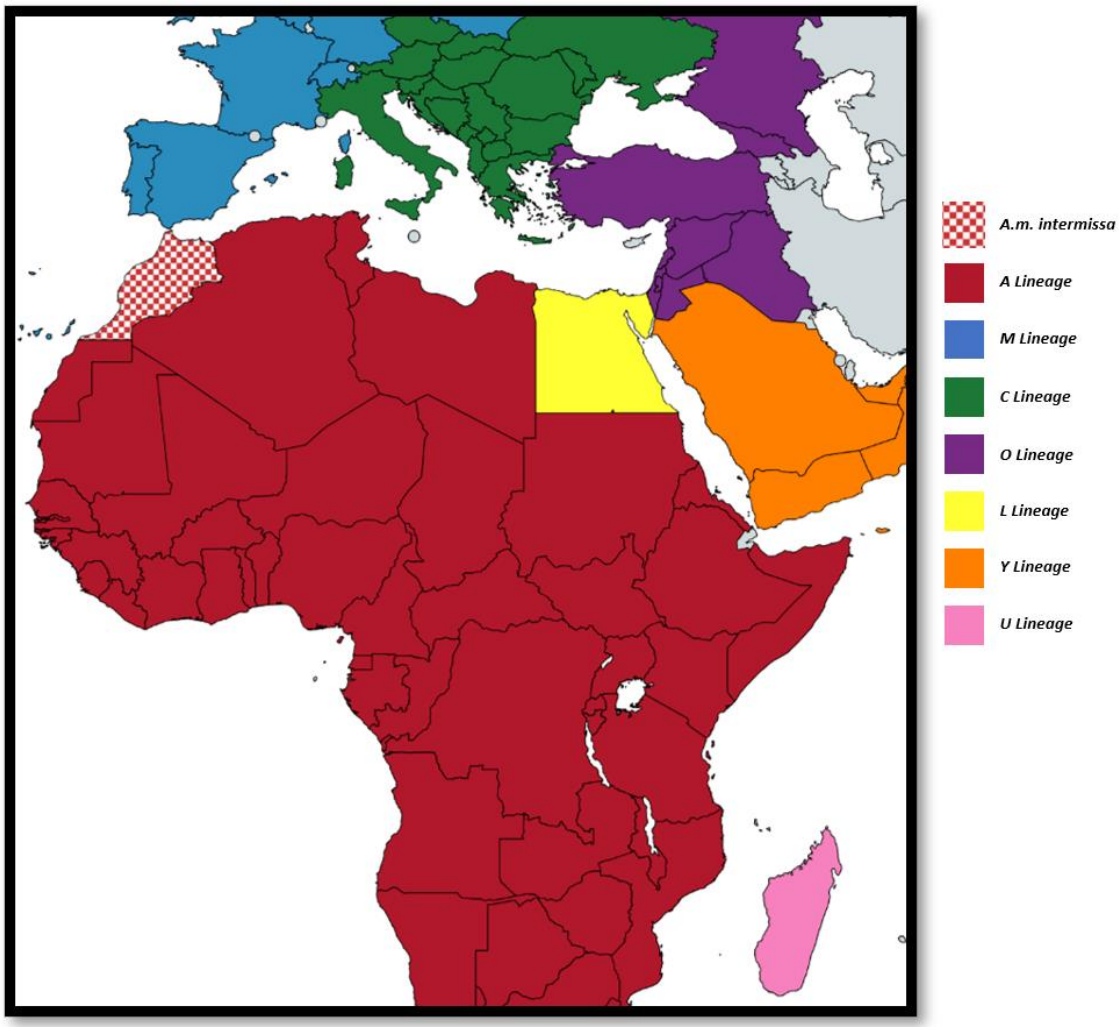

Figure 1. The geographic distribution of *Apis mellifera* throughout Europe, Africa, and Western Asia. The species is categorized into seven evolutionary lineages: M (Eurasia), C (Eastern Europe), O and Y (Western Asia), L (Egypt), U (Madagascar), and A (Africa). *Apis mellifera intermissa,* part of the A lineage, is marked with a red checker pattern in its collection location throughout Morocco.

Admixture

Admixture is the process by which previously distinct groups interbreed, producing offspring which are a blend of the two parent populations. Numerous factors can contribute to admixture events, including changes in climate or habitats, shifts in behavioral patterns such as migration, and human mediated dispersal such as historic or contemporary transportation of honey bees for apiculture. Such events erode the geographic or ecological barriers that once isolated populations, facilitating gene flow between genetically distinct groups and setting the stage for admixture (Dogantzis et al., 2021; Harpur et al., 2012; Mallet, 2005). Interspecific admixture is estimated to occur at a rate of 25% in plants and 10% in animals, while intraspecific hybridization rates vary greatly (Mallet, 2005). For example, a documented instance in the waterfowl *Anatida*, showed propensity to hybridize in its native range and in captivity at an estimated rate of 75% (Mallet, 2005; Tubaro & Lijtmaer, 2002). Comparatively, some species have significantly lower hybridization rates, such as *Drosophila*, with a reported intraspecific hybridization rate of (1/10,000) (Moehring et al., 2007; Noor, 2005). In some cases, human management appears to be a major driver of intraspecific hybridization, with most captive North American tigers possessing highly admixed genomes, and managed Canadian honey bees displaying significant admixture (Armstrong et al., 2024; Harpur et al., 2012).

A large body of evidence suggests that admixture can increase population's adaptive potential through the introduction of novel genetic material, a phenomenon known as hybrid vigor (Adavoudi & Pilot, 2021). Conversely, hybrid individuals might be less well-adapted to their environment than their parent populations, especially when admixture disrupts locally adapted genotypes. At the extreme, this can lead to the genetic identity of a locally adapted species being erased by an incoming donor population—a process known as gene swamping.

When gene swamping occurs alongside reduced hybrid fitness, it can ultimately drive a species toward extinction (Adavoudi & Pilot, 2021; Mallet, 2005).

When the genomes of two genetically distinct populations undergo hybridization, the offsprings genome will be a blend of each population. Over time, continued breeding within the population causes these population-specific segments to disperse throughout the genome. These segments, known as ancestral blocks, can be identified within the genome of a hybrid individual. Through each successive breeding event, linkage disequilibrium among the ancestral haplotype blocks slowly decay, due to gametic recombination events. Recent hybridization results in large, intact ancestral blocks, while recombination over successive generations shuffles genetic material, leading to shorter, less connected tracts of detectable ancestry. Thus, the length and maintenance of each ancestral block can be understood as a function of time since the initial admixture event and the duration of gene flow (Buerkle & Lexer, 2008; Moorjani & Hellenthal, 2023; Moorjani et al., 2011).

The identification and subsequent analysis of these blocks is known as local ancestry analysis, which seeks to locate ancestrally derived genomic regions in order to gain an understanding of how lineage-specific alleles are either maintained or lost over time. This approach not only reveals which genetic variants may offer adaptive advantages to hybrid populations but can also estimate the timing of admixture events. Understanding such estimates can illuminate the historical context of interbreeding events, aiding in the reconstruction of migration patterns and evolutionary timelines. Collectively, these techniques enhance our understanding of how genetic admixture impacts the dynamics of species evolution and adaptation (Mallet, 2005; Moorjani & Hellenthal, 2023).

Apis mellifera is an excellent model for studying admixture due to its broad distribution, which has led to unique population structure, complemented by its historically well-defined genetic lineages (Dogantzis et al., 2021; Dogantzis & Zayed, 2019). Admixture in *A. mellifera* is common, driven by the geographic distribution of subspecies and human-mediated dispersal from apiculture practices (Calfee et al., 2020; Carpenter & Harpur, 2021; Harpur et al., 2012). Furthermore, the mating behavior of *A. mellifera* is polyandrous. A virgin queen bee will mate with up to 20 drones in a single mating event, storing their genetic material to use throughout her lifetime. This behavior significantly heightens the chances of interbreeding between different lineages (Crozier & Page, 1985; Crozier & Fjerdingstad, 2001). Admixture analysis can enhance our understanding of *Apis mellifera* by aiding in the reconstruction of historical movement and hybridization patterns, revealing the ancestral contributions to key loci and the effects of hybridization on genetic diversity. These insights will illuminate the complex relationship between admixture, adaptation, and genetic diversity in one of the world's most important ecological and economic insect species.

Genetic Diversity and Admixture

Genetic diversity is essential for enabling species to adapt to changing environments by generating novel genetic combinations. A population's genetic diversity can be increased through processes like mutation, gene flow, and balancing selection, while other forces such as genetic drift, population bottlenecks, founder effects, and directional selection can reduce it (Frankham, 2002).

Admixture can offer the recipient population new genetic material, acting as a mechanism to increase genetic diversity. There are examples of Admixture increasing genetic diversity leading to notable fitness advantages in the recipient population. For instance, in *Silene*

vulgaris, a plant commonly known as Bladder Campion, admixed North American populations displayed significantly higher fitness compared to non-admixed varieties in the same environment (Keller & Taylor, 2010). A compelling example of admixture promoting genetic diversity comes from a case study of a small lizard introduced to Florida. When a species is introduced to a new location, it's generally expected that the genetic diversity of the founding individuals will be limited compared to that of the original population. However, *Anolis sagrei* experienced an unexpected increase in genetic diversity due to multiple introductions from different source populations—contrasting the usual pattern of reduced diversity seen in biological invasions (Kolbe et al., 2004). Genetic diversity is also a common concern in managed animal populations, where selective breeding can create a manufactured founder effect and lead to a decrease in genetic diversity. However, in Canadian honey bee populations, management practices have resulted in increased admixture, thereby boosting genetic diversity (Harpur et al., 2012). This is particularly important as research has shown that increased genetic diversity enhances pathogen resistance and colony fitness within honey bee colonies (Mattila & Seeley, 2007; Tarpy, 2003).

Apis mellifera intermissa

Apis mellifera intermissa, also known as the "Tellian" bee, is a North African honey bee subspecies within the A lineage, which encompasses honey bees across the African continent. Buttel-Reepen chose the name intermissa during the subspecies' first major documentation and classification in 1906, reflecting its role as an intermediate among African honey bee subspecies and a crucial link between the African (A) and European (M) lineages. Positioned between the Atlas Mountains and the Mediterranean, *A. m. intermissa* spans a vast range across North Africa, where it serves as a genetic and morphological bridge (Ruttner, 1988).

Morphometric data relieved that *A.m intermissa* acts as a connecting link between European bees (*A. m.sicula* and *A. m.iberica*) and African bees (*A. m.saharensis* and *A. m. scutellata*) when assessing physical characteristics. Among these, its closest relationship is with *A. m. sicula*, commonly known as the Sicilian bee, native to Italy. Physically, *A. m. intermissa* is known for the shiny black pigmentation of its exoskeleton, yet has a smaller body size compared to other dark bodied bees, which are primarily the European subspecies such as *A. m. mellifera* and *A. m. iberiensis* (Ruttner, 1988).

Apis mellifera intermissa is highly adapted to North Africa's seasonal climate, marked by wet winters and dry summers. Floral resource availability is heavily influenced by fluctuations in rainfall, requiring *A. m. intermissa* to time its resource-dependent tasks, such as colony reproduction, to coincide with windows of sufficient nectar availability. *A. m. intermissa* experiences two seasonal reproductive peaks; this reproductive pattern, known as a bimodal brood cycle, is synchronized with the Mediterranean floral cycle. In spring, abundant nectar supports active reproduction and feeding. The summer drought causes nectar availability to drop, which is followed by a decrease in swarm production rates in *A. m. intermissa* populations. A secondary brood peak occurs from fall to winter during moderate floral resource supply. This bimodal brood rhythm enables the species to thrive despite periods of food scarcity. Interestingly, *A. m. intermissa* possesses one of the highest reproductive capacities of all *Apis mellifera* subspecies. During the height of its reproductive season, a colony can produce over 100 queen cells. These cells give rise to new queens, which, upon reaching maturity, initiate the formation of new honey bee colonies through a process called swarming. This adaptation allows *A. m. intermissa* to rapidly recover from colony losses during periods of low productivity (Ruttner, 1988).

Beekeepers have noted some of the more challenging behaviors of *Apis mellifera intermissa*, including a strong defensive response, high nervousness, and heavy use of propolis. However, *A. m. intermissa* exhibits a wide range of behavioral variation across its range, making it challenging to describe definitive characteristics of the subspecies (Ruttner, 1988). Despite criticisms regarding its temperament, *A. m. intermissa* exhibits a remarkable trait of great economic importance: increased resistance to the brood parasite *Varroa destructor*. This resistance is attributed to its effective grooming behavior (Adjlane et al., 2016; Boecking & Ritter, 1993; Dadoun et al., 2020; Locke, 2016), which provides a significant advantage, as *Varroa destructor* is a major factor in colony health challenges affecting *Apis mellifera* populations globally (Noël et al., 2020).

Microsatellite and mitochondrial data research has revealed foundational information regarding the genetic relationships and population structure of *Apis mellifera intermissa*. According to research by Abou-Shaara et al. (2021), *A. m. intermissa* populations in Morocco and Algeria possess significantly different genetic compositions, with Moroccan *A. m. intermissa* populations sharing a closer relationship to Egyptian bees (*A. m. lamarckii*), while Algerian populations are more genetically similar to the African subspecies *A. m. scutellata* (Abou-Shaara et al., 2021). Furthermore, these genomic techniques have revealed ongoing hybridization between *A. m. sahariensis* and *A. m. intermissa* in Moroccan populations. This hybridization raises concerns about the genetic purity of subspecies across north Africa *(Aglagane et al., 2023; Chahbar et al., 2013).*

Apis mellifera intermissa was analyzed using whole-genome SNP data in multiple studies. In a landmark phylogenetic study by Whitfield et al. (2006), 14 *Apis mellifera* subspecies were analyzed at 1,136 genomic SNPs, including *Apis mellifera intermissa*, to

identify the most probable adaptive radiation pattern of *Apis mellifera*. The results of this work supported an African origin for *Apis mellifera*. This conclusion was later challenged by Dogantzis et al. (2021), who expanded the dataset to 11.8 million SNPs and proposed an Asian origin instead. Interestingly, the Moroccan *A. m. intermissa* population used in the work of Whitfield et al. (2006) was re-sequenced in the study by Dogantzis et al. (2021), but was excluded from the major phylogenetic analysis due to its admixed genetic composition.

Despite these differing conclusions, both studies consistently identified the Moroccan *A. m. intermissa* as an admixed population, with ancestry proportions of approximately 73% African and 27% European. This genetic composition has been central to debates regarding the M lineage's origins, a hypothesis that remains contested (Abou-Shaara et al., 2021; Carr, 2023; Cridland et al., 2017; Dogantzis et al., 2021; Whitfield et al., 2006). While its admixed nature complicates phylogenetic analyses, *A. m. intermissa* provides a valuable model for studying admixture dynamics. Further investigation into its genomic structure could help clarify honey bee admixture patterns in North Africa and their broader evolutionary significance.

Admixed populations serve as natural laboratories for evolutionary biology, offering insights into the genomic outcomes of mixing distinct lineages (Hewitt, 1988). Estimating the timing of admixture events can illuminate honey bee dispersal patterns, while local ancestry analysis can pinpoint lineage-specific genetic contributions. These contributions may include traits linked to resilience when mapped to specific genes (Moorjani & Hellenthal, 2023). Despite its potential, the admixture patterns within this unique *A. m. intermissa* population remain underexplored, limiting our understanding of how admixture influences population structure, genomic diversity, and adaptability. Addressing this gap could enhance conservation strategies,

inform breeding programs, and expand our understanding of the ecological and evolutionary impacts of genetic admixture.

To characterize the admixture patterns in *Apis mellifera intermissa*, we analyzed 15 genomes sampled from populations throughout Morocco where admixture has been previously identified (Dogantzis et al., 2021; Whitfield et al., 2006). First, we aimed to assess the population structure and ancestral composition of the admixed *Apis mellifera intermissa* population, and to visualize it within the broader context of *Apis mellifera* lineages and subspecies. To do this, we utilized the maximum likelihood estimation approach as implemented in ADMIXTURE, complemented by principal component analysis (PCA) to delineate similarities in genetic composition across the major *Apis mellifera* lineages. (D. H. Alexander et al., 2009).

Second, by comparing the lengths of African (A lineage) and European (M lineage) ancestry tracks in the *Apis mellifera intermissa* population, we estimated the number of generations that have passed since the admixture event using a method implemented in ROLLOFF (Moorjani et al., 2011). Given that the M lineage likely spread into Europe via a northern route out of Asia, North Africa is an improbable historical contact zone, suggesting that admixture in this region may stem from more recent secondary contact (Dogantzis et al., 2021). Additionally, M-lineage *A. m. iberiensis* populations adjacent to *A. m. intermissa* in Spain exhibit A-lineage admixture, reflecting recent secondary contact (Chávez-Galarza et al., 2015). These findings support our hypothesis that the admixture observed in *A. m. intermissa* may similarly reflect recent admixture, rather than admixture from a historical contact zone.

By analyzing *A. m. intermissa* and comparing it with the African (A) and European (M) lineages, we assessed diversity metrics, including nucleotide diversity (π) and expected/observed heterozygosity, to characterize genetic diversity in both the progenitor populations and the admixed subspecies.

Finally, we conducted local ancestry analysis by pinpointing stretches of the genome inherited from ancestral sources, identifying regions with significant European (M lineage) genetic contribution using a method implemented in LOTER (Dias-Alves et al., 2018).

Methods

Sample Curation

A total of 15 whole genome sequences of *Apis mellifera intermissa* were retrieved from the study by Dogantzis et al. (2021), which re-sequenced the *A. m. intermissa* samples originally collected and published in the work of Whitfield et al. (2007). This population was sampled across Morocco, with each worker bee collected from a different hive. To support the analysis, a library of 119 *A. mellifera* genomes were retrieved from the works of Dogantzis et al 2021, Wallberg et al 2017, and Harpur et al 2014, representing the A, C, O, M, L and Y lineages (Supplementary Table S1). These genomes were distributed among the known evolutionary lineages as follows: A Lineage: 60 samples, M Lineage: 18 samples, C Lineage: 12 samples, O Lineage: 11 samples, Y Lineage: 10 samples, and L Lineage: 8 samples.

Sequence Processing of Sample Genomes

Sequenced genomes were processed using Trimmomatic- 0.40 (Bolger et al., 2014) to remove Illumina sequencing Adapters, and low-quality sequence reads. Reads were trimmed if they had a Phred score below 20, a threshold ensuring 99% base call accuracy. After trimming,

the quality of the trimmed reads was assessed using FASTQC-0.12.0 ("FastQC," 2015). Next, trimmed samples were aligned to the Amel_HAv3.1 (Wallberg et al., 2019) genome assembly. This alignment process was performed using the NextGen Map-0.5.0 (Sedlazeck et al., 2013) aligner. The quality of the alignments was evaluated using QualiMap-2 (Okonechnikov et al., 2015). This tool assesses the accuracy of the alignment positions. As a quality control measure, only samples with at least 95% of reads successfully mapped were retained for analysis. Additionally, the samples were evaluated based on several quality metrics, including duplication rate, mean mapping quality, mean coverage, and GC content. Next samples were sorted using SAMtools-1.3.1 (Danecek et al., 2021) to ensure that all sequencing reads are arranged in order according to their positions on the reference genome. Duplicate reads, which may arise from PCR amplification or sequencing artifacts, were identified, and flagged using Picard toolkit (*Picard Toolkit*, 2019). Next, to optimize the storage and retrieval processes, an index of the sequencing data was created using SAMtools-1.3.1(Danecek et al., 2021). For the identification of single nucleotide polymorphisms (SNPs) within the samples, BCFtools (Danecek et al., 2021) was utilized. Variant calling detected 27,456,741 SNPs.

The genomes of eight haploid drones from Dogantzis et al. (2021) were mapped to the AmelHav3.1 reference genome (Wallberg et al., 2019), and treated as diploid samples, following the same protocol as the previously mentioned worker samples. During the filtering process, regions displaying heterozygous base calls were flagged. These regions signify uncertainty due to extensive repeats and low-quality base calls. We excluded mutations within a 10 bp window of any SNP what was heterozygous in our haploid drone dataset from our analysis. Additionally, SNPs located within 10 base pairs of an insertion or deletion (indel) were removed to reduce the likelihood of false positives.

I further filtered SNPs based on depth, Mapping Quality Strand Bias Z-score (MQSBZ), and Read Position Bias Z-score (RPBZ). Values were excluded using an IQR outlier detection method, to remove values deemed excessively high or low, which are indicative of low-quality SNPs (Supplementary Figure S1). I applied specific thresholds using BCFtools filter (Danecek et al., 2021) with the following conditions: depth between 5,642 and 11,322, MQSBZ between -1.204642 and 1.216214, and RPBZ between -3.998202 and 4.047033. Additionally, SNPs were excluded if the Mapping Quality Zero Fraction (MQ0F) metric, which indicates the proportion of zero-quality mappings, exceeded 0.0921153, discarding the top 5% of values associated with low-quality mappings (Supplementary Figure S1). SNPs were also excluded if they had 10% or more missing genotype values. Following these SNP filtering criteria, 16,510,464 SNP sites remained for subsequent analyses.

Population Structure

To evaluate the population structure of the *Apis mellifera* samples, I performed a principal component analysis (PCA) and an admixture analysis. To reduce the effects of linkage disequilibrium, SNPs were pruned to a distance of 5000bp prior to analysis (Wallberg et al., 2015). Additionally, the SNP set was pruned of low frequency variants by filtering loci with a minor allele frequency (MAF) less than 0.05 (Linck & Battey, 2019). This filtering process resulted in 41095 biallelic SNPS. The curated dataset was subsequently used to analyze population structure, starting with global admixture analysis of the samples with ADMIXTURE v1.3.0 (David H Alexander et al., 2009). The analysis was run with predicted K values 3-10 and included the 10-fold cross-validation procedure to estimate the optimal value for *K*. Then a PCA was performed to further clarify unique genetic clusters within the dataset. This analysis was conducted in R-4.2.1 using SNPRelate (Zheng et al., 2017; Zheng et al., 2012).

Genetic Diversity: π Heterozygosity

For calculating diversity metrics, the raw SNP dataset was used without applying a minor allele frequency (MAF) threshold of < 0.05. Avoiding this threshold prevents the exclusion of rare variants, which could potentially bias the calculation (Everitt et al., 2023; Eynard et al., 2015; Schmidt et al., 2021). Nucleotide diversity (π) was estimated using VCFtools (0.1.15) (Danecek et al., 2011), applying 500 bp sliding windows with a 250 bp step size to provide fine-scale resolution. This analysis was conducted on the progenitor A (n = 20) lineage, the M lineage (n = 18), and the *A. m. intermissa* (n = 15) subspecies, to make comparative assessments between hybrid and founding population diversity. In addition, observed and expected heterozygosity values were calculated for each SNP, across the progenitor lineages A and M, and the admixed group *A. m intermissa*. These heterozygosity metrics were averaged within the same 500 bp windows used for nucleotide diversity to ensure methodological consistency across analyses. To minimize the effects of linkage disequilibrium and ensure independence among observations, all diversity metrics were subsequently averaged into non-overlapping 5,000 bp genomic blocks. A custom script was developed to perform these calculations in R-4.4.2, publicly available at (https://github.com/NotoriousBEE/Heterozygosity_Genome/blob/e83ad0bff1418c823b62bc8d1f5f5f0f035e5de0/Heterozygosity_windows). Due to non-normality observed in the diversity distributions, statistical differences among lineages were assessed using the non-parametric Friedman test, followed by pairwise Wilcoxon signed-rank tests with Bonferroni correction for post-hoc comparisons.

Local Ancestry

I estimate local ancestry across the genome of 15 *A. m. intermissa* using 20 samples of *A. m. scutellata* (A lineage) and 18 samples from the M lineage (representing all available M lineage samples in the dataset), as reference samples. To reduce redundant SNPs and optimize the analysis, the filtered SNP dataset was pruned for ancestry-informative markers (AIMs). SNPs were classified as AIMs if their Fst value, calculated between progenitor lineages A and M, were in the top 25th percentile (Fst $\geq$ 0.069), resulting in a final set of 4,336,281 SNPs. The samples were phased and missing data was imputed using Beagle5.1 (Browning et al., 2021; Browning et al., 2018). Local ancestry inference was performed on the *A. m. intermissa* population using LOTER (Dias-Alves et al., 2018) with bagging enabled, with the A and M populations used as source populations. M lineage genomic enrichment was assessed by identifying loci ranked in the 99th percentile of M lineage frequency at a given SNP site, which consecutively spanned a region exceeding the expected lineage threshold (length > 5kb).

Time Since Admixture

The variant set, enriched for ancestry-informative markers (AIMs), was used to estimate the number of generations since admixture. We used a previously published 100 bp-scale linkage map for *Apis mellifera* from Jones et al. (2019). The map was then converted from the population recombination rate (ρ or rho) to the recombination rate (r) using the formula $\rho = 3Ner$, where Ne represents the effective population size, and the factor of 3 accounts for the haplodiploid nature of honey bees (Wallberg et al., 2015). An effective population size of 442,757, estimated for the A lineage, was used for the conversion (Jones et al., 2019). This value was chosen because the analysis targeted hybrid *A. m. intermissa* bees, which are predominantly A lineage in genetic composition. This analysis was performed using the ROLLOFF method (Moorjani et al., 2011).

In this approach, LD decay within the admixed population was calculated and then fitted to an exponential distribution using least-squares regression, which allows for the estimation of the number of generations that have passed since the admixture event (Moorjani et al., 2011).

Functional Enrichment Analysis and QTL Enrichment

To determine the possible functional contributions of the enriched M lineage genes in this *A. m. intermissa* population, additional analyses were performed. Functional enrichment analysis and QTL enrichment assessment were used to identify biological processes associated with enriched M lineage segments.

Functional annotation cluster analysis was identified using DAVID v6.8 (Huang et al., 2009; Sherman et al., 2022) utilizing genes from BEEBASE (Walsh et al., 2022). The SNPs utilized as background genes were those pruned for Ancestrally Informative Markers (AIMs) in the local ancestry analysis, specifically those that intersected genes within the *Apis mellifera* reference genome's (AMEL_HAv3.1) gnomic feature file. This process identified 8,409 unique background genes for the functional annotation cluster analysis. Additionally, genes that fell within M lineage enriched regions were identified, resulting in 166 unique target genes. Clusters were determined to be meaningful if they had an enrichment score ≥ 2 (Huang et al., 2009; Sherman et al., 2022).

Various QTLs have been identified that are of interest in hybrid A and M populations, because these regions contribute to defense response, varroa sensitive hygiene, hygienic behavior, and grooming (Calfee et al., 2020). The QTL segments of interest were mapped to the current *A. mellifera* HAv3.1 genome (Wallberg et al., 2019) and compared to enriched M lineage regions in the *A. m. intermissa* population to assess M lineage genetic contribution to traits influenced by the QTLs (Calfee et al., 2020).

Results

Population Structure

I estimated the global ancestry proportions of *Apis mellifera intermissa*, Using several population structure analyses. The ADMIXTURE 10x cross validation results suggest the optimal number of genetic clusters to be K = 7 (Supplementary Figure S3), but we find that K = 5 better aligns with the biological context of the populations analyzed. When K = 5, native *Apis mellifera* samples cluster into previously identified lineages (Figure 2). The exception is the L and A lineages that form a single genetic cluster, likely due to a shared recent common ancestor and close evolutionary history (Figure 2) (Dogantzis et al., 2021). At higher K values (K = 6–8), *A. m. intermissa* consistently forms its own distinct cluster, reaching ~99% assignment to a single ancestry component, reflecting its genetic distinctiveness despite its previously established admixed origins (Supplementary Figure S2). Notably, at K = 7, *A. m. lamarkii* (L lineage) further separates from the broader African lineage. These additional subdivisions, while informative, introduce finer-scale structure that complicates interpretation. In contrast, K = 5 captures the major lineage splits while preserving a biologically interpretable structure consistent with prior literature and the known evolutionary history of this *A. m. intermissa* population (Dogantzis et al., 2021; Whitfield et al., 2006). The results of the global admixture inference highlight that samples of *A. m. intermissa* are highly admixed, composed primary of A lineage ancestry (74.1% +/- SE 0.00616), with the remainder originating from the M lineage (25.9% +/- SE 0.00616). These patterns are congruent with previous studies (Dogantzis et al., 2021; Han et al., 2012b; Whitfield et al., 2006). The results of the PCA reflect the patterns found with the ADMIXTURE analysis (Figure 3). The first two principal components of the PCA captured 23.88% and 21.59% of the genetic variance in the dataset (Figure 3). Representative honey bee samples cluster into

the six main lineages, while *A. m. intermissa* is positioned centrally between the major clusters representing the A and M lineages, emphasizing the admixed origins of the subspecies.

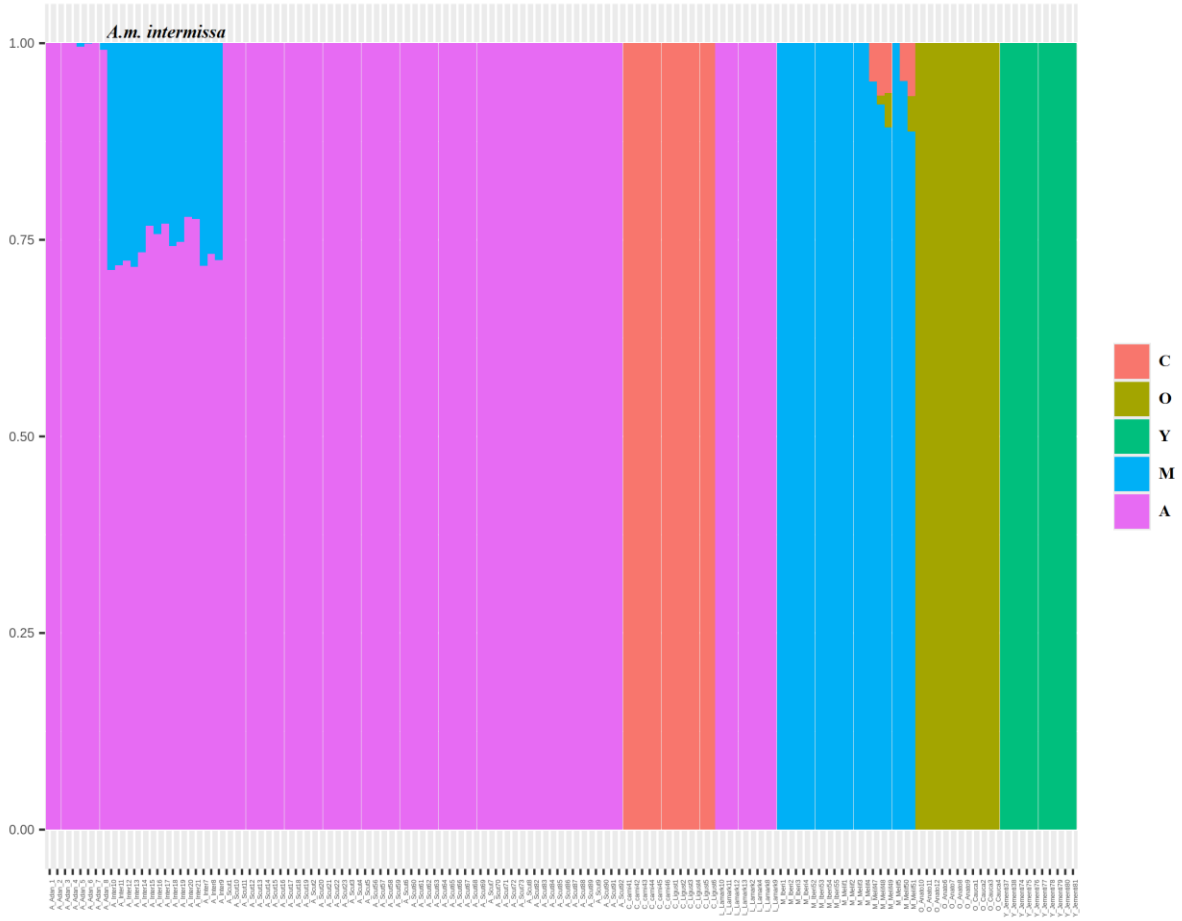

Figure 2. Ancestry proportions of *Apis mellifera* lineages inferred through global admixture analysis of N=134 Apis mellifera samples with K=5 genetic clusters. Each vertical bar represents an individual bee, with colors indicating the proportion of its genome attributed to each ancestral population.

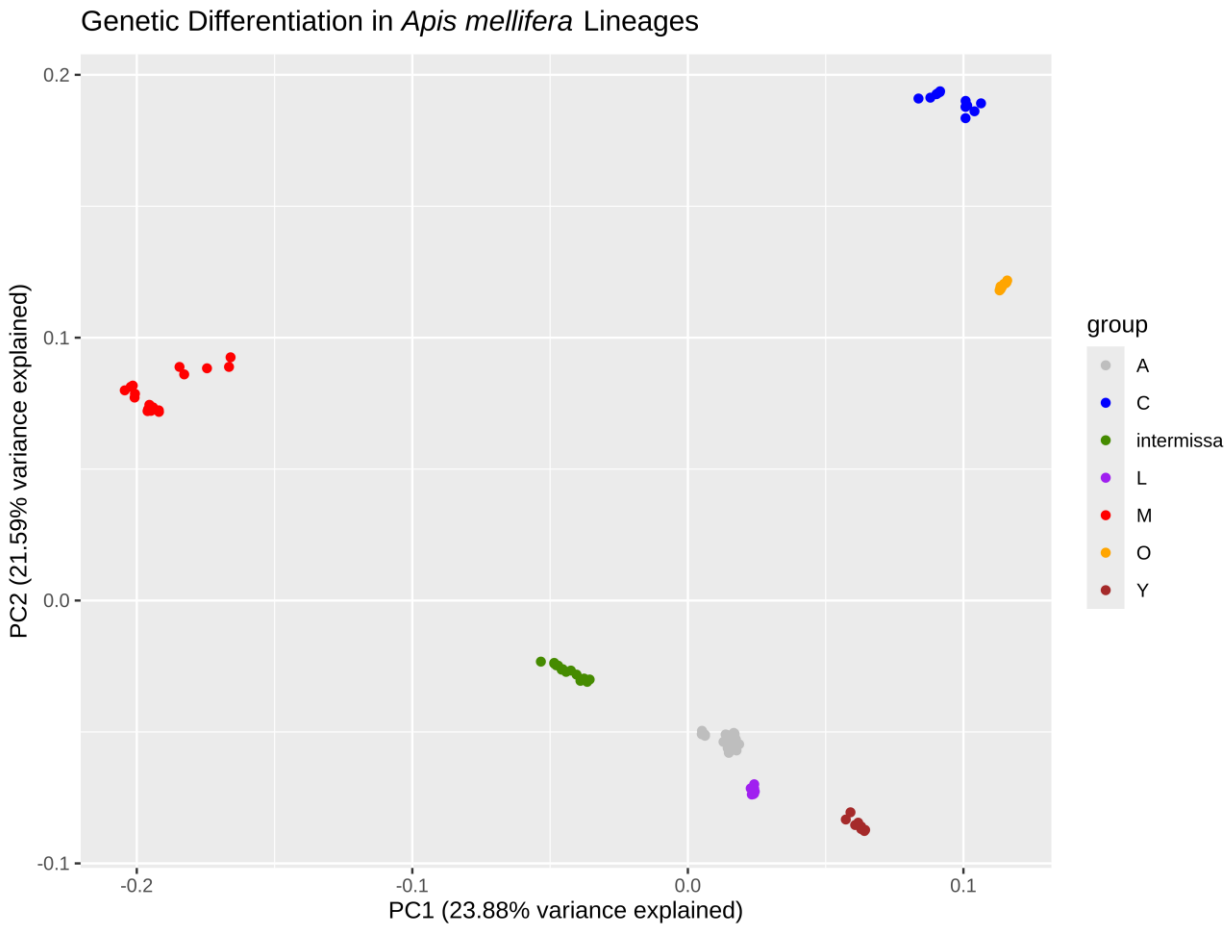

Figure 3. Principal component analysis (PCA) reveals that *A. m. intermissa* is genetically intermediate between the African (A) and European (M) honey bee lineages. The PCA was constructed using *Apis mellifera* whole-genome SNPs with a minor allele frequency (MAF) > 0.05, focusing on genomic variants to emphasize lineage-level population structure and patterns of admixture. The plot shows seven groupings representing six *A. mellifera* lineages and the admixed subspecies *A. m. intermissa,* positioned between the A and M lineages.

Genetic Diversity: π and Heterozygosity

I assessed the genetic diversity in the progenitor lineages A and M, as well as the admixed lineage *A. m. intermissa*, by measuring nucleotide diversity (π). The results indicate that lineage A has the highest genome-wide nucleotide diversity ($\pi = 0.00391$), while lineage M exhibits the lowest diversity ($\pi = 0.00119$), consistent with previous studies (Cridland et al., 2017; Dogantzis et al., 2021). The admixed lineage *A. m. intermissa* displays an intermediate level of nucleotide diversity ($\pi = 0.00360$), falling between the values observed for the A and M lineages. Statistical analysis using A non-parametric Friedman test confirmed that differences in nucleotide diversity among the three groups were highly significant ($\chi^2 = 752{,}328$, $df = 2$, $p < 0.001$; Supplementary Figure S4). A post-hoc pairwise Wilcoxon signed-rank test revealed significant differences between all groups, with lineage A exhibiting significantly higher nucleotide diversity than lineage M ($V = 331{,}540{,}541{,}903$, $p < 0.001$), *A. m. intermissa* showing significantly higher diversity than lineage M ($V = 15{,}166{,}757{,}538.5$, $p < 0.001$), and *A. m. intermissa* displaying lower diversity than lineage A ($V = 216{,}282{,}407{,}061.5$, $p < 0.001$) (Supplementary Figure S4).

Heterozygosity calculations mirror the trend observed in nucleotide diversity, with both expected and observed heterozygosity values being highest in the A lineage (expected = 0.0548; observed = 0.0542), followed by the admixed *A. m. intermissa* group (expected = 0.0460; observed = 0.0465), and the lowest mean values recorded in the M lineage (expected = 0.0223; observed = 0.0119). The Friedman tests for both expected ($\chi^2 = 1{,}454{,}544$, $df = 2$, $p < 0.001$) and observed heterozygosity ($\chi^2 = 1{,}395{,}652$, $df = 2$, $p < 0.001$) revealed highly significant differences among populations (Supplementary Figures S5–S6). Post-hoc pairwise Wilcoxon signed-rank tests with Bonferroni correction indicated that lineage A had significantly higher

heterozygosity values compared to lineage M (expected: V = 347,584,016,410.5, p < 0.001; observed: V = 347,551,619,906.0, p < 0.001). The admixed lineage *A. m. intermissa* exhibited intermediate heterozygosity (expected: V = 347,584,016,410.5, p < 0.001; observed: V = 347,551,619,906.0, p < 0.001), and significantly higher than lineage M (expected: V = 1,257,047,451.0, p < 0.001; observed: V = 1,010,110,142.0, p < 0.001) and significantly lower than lineage A to (expected: V = 328,346,783,494.0, p < 0.001; observed: V = 313,069,064,297.0, p < 0.001) (Supplementary Figures S5–S6).

Table 1. *A. m. intermissa* exhibits intermediate genomic diversity between its progenitor lineages A and M**.** Key metrics include nucleotide diversity (π) and expected and observed heterozygosity. Sample sizes were balanced, with the M lineage incorporating all subspecies groups to increase representation. The reported number of SNPs represents sites where at least one copy of the variant allele was present in the population.

Subspecies	Population	No. Genomes	No. SNPs	(π)	Expected Heterozygosity	Observed Heterozygosity
A. m. intermissa	***A. m. intermissa***	15	16313277	0.0036	0.0460	0.0465
A. m. scutelata	**A**	20	16348484	0.0039	0.0548	0.0542
A. m. mellifera + *A. m. Iberiensis*	**M**	18	16269096	0.0012	0.0223	0.0199

Identifying Regions Enriched for M Ancestry

Local ancestry was estimated using ancestrally informative markers (AIMs) for the A and *M* lineages. We then assessed whether specific genomic regions were enriched for ancestry from the minor donor lineage (*M*), potentially conferring a genetic advantage over the dominant *A*-lineage. The local ancestry analysis obtained a slightly higher estimation of M lineage contribution at 33.9% (SD ± 0.0876), compared to the estimate obtained during global ancestry

analysis at 25.9%. We further analyzed the results for *M* lineage enrichment, identifying loci within the top 1% of *M* ancestry (≥80% M ancestry at a SNP). Using this threshold, we discovered 89 genomic intervals enriched for the *M* lineage in the *A. m. intermissa* population (Figure 4). Within the intervals, 10,343 SNPs were fixed with M lineage contribution. The average interval length was 28,200 bp, with the longest segment measuring 412,423 bp on chromosome 11 (Fig. 4). The largest number of enriched intervals were located on chromosome 7, which comprised of 25.84% of all enriched intervals (Figure 4).

Time Since Admixture

To estimate the number of generations that occurred following a single pulse admixture model, ROLLOFF (Moorjani et al., 2011) was used, which examines the decrease in admixture linkage disequilibrium between source and hybrid populations. The African (n = 20) and European (n = 18) lineage groups were selected to represent ancestral source populations. The number of generations since admixture was estimated by fitting an exponential distribution to the ROLLOFF (Moorjani et al., 2011) calculations of linkage disequilibrium decay. In the Moroccan *A. m.intermissa* population, admixture was estimated to have occurred 14.04 generations ago (*SE = 0.325,* Supplementary Figure S7).

Characterization of Candidate Genes

We conducted a comprehensive functional enrichment analysis using DAVID (Database for Annotation, Visualization, and Integrated Discovery) v2024q4 (Huang et al., 2009; Sherman et al., 2022), specifically focusing on gene clusters to explore whether genomic regions enriched for M ancestry are associated with broad functional features. By leveraging all available annotation databases—including Gene Ontology (GO) terms, KEGG pathways, and protein domains—we aimed to identify enriched biological functions, pathways, and structural features

within the clustered genes in our target list. One suggestive annotation cluster was identified, achieving an enrichment score of 2.23. This cluster comprises genes associated with Juvenile hormone production, which influences age differentiation and division of labor (Elekonich et al., 2001). This cluster contains notable terms such as farnesyl diphosphate biosynthetic process (P value 2.3E-3, Benjamini 4.6E-1), and dimethylallyltranstransferase activity (P value 2.5E-3, Benjamini 1.2E-1). The cluster consists of the genes GB46096, GB48898, GB48897. (Supplementary Table S2).

Enriched Regions and Associated QTLs

Of the 89 regions enriched for M lineage within the *A. m. intermissa* population, 18 segments are placed within 4 quantitative trait loci's (QTLs) associated with Varroa Sensitive Hygiene (Figure 4) (Spötter et al., 2012). Eleven enriched segments overlapped with a QTL on chromosome 1 and two enriched segments coincided with a QTL on chromosome 2. The QTL on chromosome 12 contained four enriched segments, and one segment overlapped with a QTL on chromosome 16 (Figure 4).

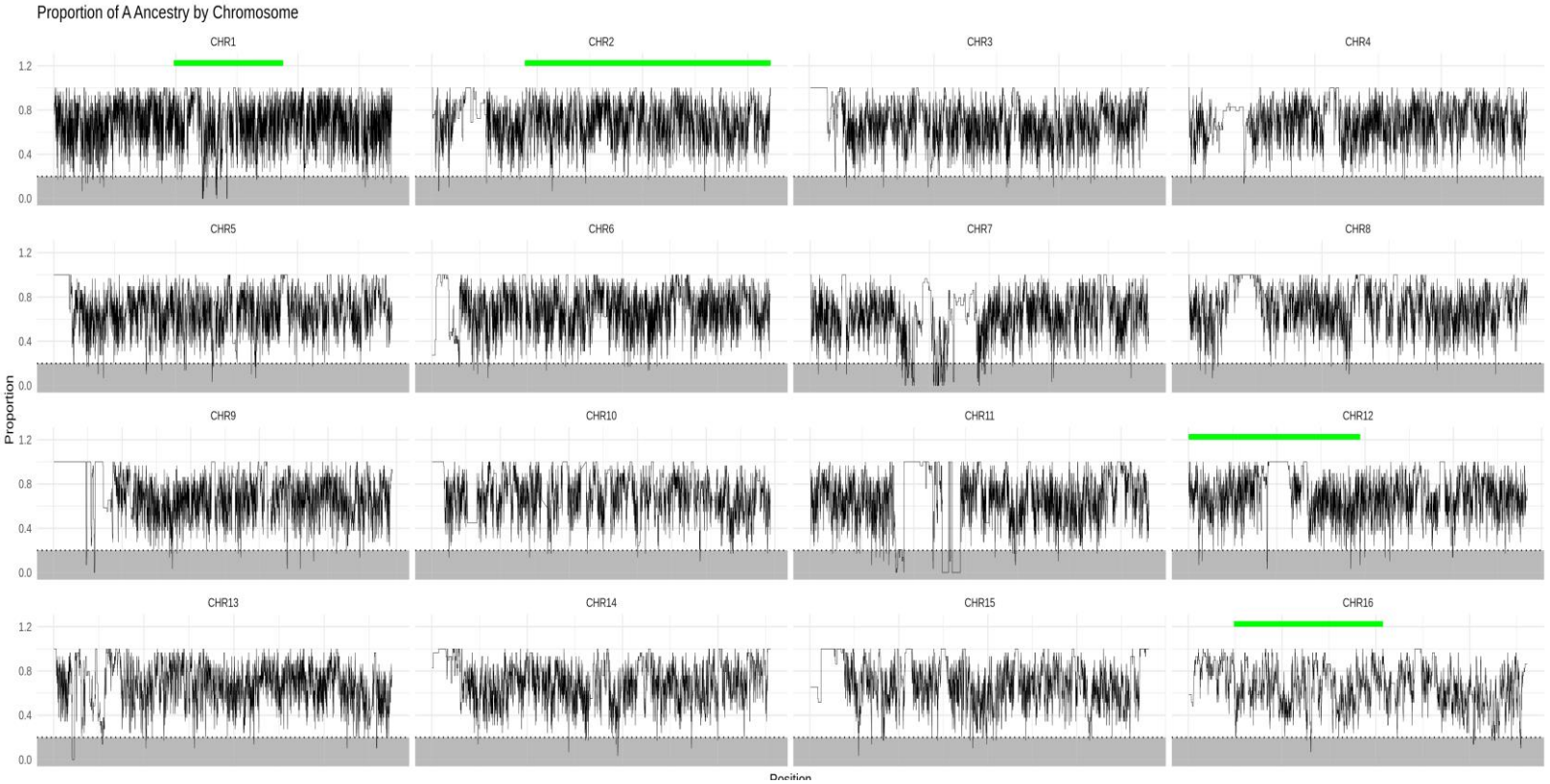

Figure 4. Proportion of A lineage ancestry across the Moroccan *A. m. intermissa* population across each chromosome. The grey shaded area illustrates the threshold where A lineage ancestry falls below the 99th percentile (20% A lineage ancestry). Green bars indicate the positions of Quantitative Trait Loci (QTLs) associated with Varroa Sensitive Hygiene, that share overlap with enriched segments. The analysis included 20 samples of A lineage, 18 samples from the M lineage, and 15 samples of *A. m. intermissa.* Local ancestry inference was performed using LOTER, with subsequent identification of genomic intervals significantly enriched for M lineage contributions.

Discussion

African (A lineage) and European (M lineage) hybrid honey bee populations have been of great interest in honey bee research over the past couple of decades, with numerous studies documenting admixture patterns across various environments (Chávez-Galarza et al., 2015; Everitt et al., 2023; Nelson et al., 2017). Among these A and M lineage admixed populations, the Moroccan *A. m. intermissa* population central to this study is particularly notable for its role in historical debates surrounding the origin and dispersal routes of *A. mellifera*.

Since its initial characterization, *A. m. intermissa* has been described as a morphological intermediate between African and European honey bees, a classification that influenced its name, inspired by the word *intermediate*. Early genetic studies reinforced this view. When the first major phylogenetic analysis of *Apis mellifera* was conducted using Moroccan *A. m. intermissa* samples from this study, the results supported the hypothesis that the M lineage emerged via a northern expansion route from Africa (Ruttner, 1988). Its historical description as a "missing link," combined with the chance sampling of a hybrid population in Morocco, further supported the idea that the M lineage dispersed around the northwestern tip of the continent (Ruttner, 1988; Whitfield et al., 2006). This hypothesis that the M lineage was derived from Africa was later refuted by Dogantzis et al. (2021), who utilized whole-genome sequences from 251 individuals representing 18 putatively identified subspecies and analyzed 11.8 million SNPs across honey bee lineages. Their findings demonstrated that the M lineage originated from a northern expansion route out of Asia. If the M lineage arose from Asia, it is clear that *A. m. intermissa* is not the stepping stone between African and European bees as once hypothesized.

Since the identification of this hybrid population, further studies explicitly investigating hybrid characteristics in *A. m. intermissa* have been limited, even though significant research has

been conducted on hybrid *A. mellifera* populations in regions such as North and South America and Spain (*Calfee et al., 2020; Harpur et al., 2012; Nelson et al., 2017; Zárate et al., 2022*). To address this research gap, the goal of this study was to characterize the genomic admixture patterns in *A. m. intermissa* by analyzing global and local ancestry, estimating the timeframe of hybridization, assessing the impact of admixture on genetic diversity, and exploring conserved genes within genomic regions enriched for European ancestry.

The first step was to validate the admixed nature of the samples, ensuring the integrity of the dataset for downstream analyses. A PCA of the unlinked SNP dataset revealed seven distinct genetic clusters, six of which corresponded to the established *A. mellifera* genetic lineages: M lineage (Western Europe), C lineage (Eastern Europe), O and Y lineages (Western Asia), L lineage (Egypt), and A lineage (Africa) (Figure 3). As expected, the admixed *A. m. intermissa* samples were positioned between the A and M lineages, consistent with their previously identified admixed status. Global ancestry proportions were subsequently estimated using ADMIXTURE, confirming that the *A. m. intermissa* samples exhibited an average of 25.9% M lineage introgression and 74.1% A lineage introgression (Figure 2). These results align with previous findings from *Dogantzis et al., 2021*, further corroborating the historical categorization of *A. m. intermissa* as an intermediary between African and European bees. Morphological analyses originally suggested this placement, a conclusion later reinforced by genetic evidence. However, subspecies-level genomic analyses have detected that some Moroccan *A. m. intermissa* populations exhibit significant admixture with *A. m. sahariensis*, another A lineage subspecies native to the Morocco and Algeria. At the same time, pure populations of *A. m. intermissa* have also been identified, highlighting the complex population structure of this subspecies *(Abou-Shaara et al., 2021; Aglagane et al., 2023; Chahbar et al., 2013)*. The aim of this analysis was

to explore the lineage-level population structure of the admixed *A. m. intermissa* population, and therefore, this subspecies-level population structure cannot be identified in this work. Further work exploring the unique population structures of North African honey bee subspecies could be valuable to shed light on locally adapted subspecies and support their conservation efforts.

The choice for K = 5 in this dataset resulted in *A. m. lamarkii*, the Egyptian honey bee (L lineage), clustering with the broader African A lineage (Figure 1; Supplemental Figure S2). While *A. m. lamarkii* separated from the broader African lineage at K = 7, its placement as a distinct lineage may require broader representation of African subpopulations to enhance lineage clarity. In the study by Dogantzis et al. (2021), *A. m. lamarkii* was resolved as the L lineage, using 18 native subspecies of *Apis mellifera.* In contrast, the current study excluded *A. m. monticola*, a highly distinct A lineage subspecies native to East Africa (Kenya and Tanzania), due to its geographic distance from *A. m. intermissa* and the already extensive representation of A lineage *Apis mellifera* in the dataset. (Dogantzis et al., 2021). These choices reflect the specific aims of the study and reinforce the importance of reference panel composition when interpreting population structure.

To further characterize the nature of this admixture, the estimated timing of the hybridization event was investigated using ROLLOFF. The results suggest that admixture in the Moroccan *A. m. intermissa* population occurred approximately 14.04 generations ago (*Supplemental Figure S7*). Given that African honey bees undergo between 6 and 12 swarms per year (McNally & Schneider, 1992), but not all swarms are successful—with only about 4% of queens returning from mating flights in North Africa due to high levels of predation (Hepburn & Radloff, 2013) —this corresponds to an effective generation time of approximately 2.08 to 4.17 years. Multiplying 14.04 generations by this range suggests that the hybridization event occurred

roughly between 29 and 58 years ago. The timing of this hybridization event, occurring within the last few decades, implies that it is likely influenced by human-mediated activities. Historically, the last period of natural access between Morocco and Europe occurred during the Last Glacial Maximum, approximately 26,000 to 19,000 years ago Bowen (2009). The recent timeframe of admixture points to anthropogenic factors, potentially linked to modern beekeeping practices that involve the international trade and transportation of bee colonies.

Similarly, in a study examining an African-European hybrid population in Brazil, Nelson et al. (2017) estimated a generation time of approximately 2 years based on a combination of hybridization time estimates and documented records of African honey bee introductions. Further research is needed to fully understand the swarming and reproductive patterns of admixed African-European hybrid *Apis mellifera* populations in and surrounding regions.

Next genome wide diversity was measured to compare diversity outcomes in hybrid and progenitor populations, diversity was measured using nucleotide diversity and by assessing heterozygosity metrics. Hybridization between a very diverse lineage, A, and a lineage with comparatively low diversity due to a historical population bottle neck, M, resulted in *A. m. intermissa* displaying a genome wide average value of π between the two progenitor population values (Table 1) (Dogantzis et al., 2021). Observed and expected heterozygosity values follow a similar pattern as the estimate of π, with *A. m. intermissa's* value between the largest A value and smallest M value (Table 1). This highlights that admixture provides an avenue to both increase and decrease the genetic diversity of a population, relative to the progenitor populations.

Although the hybridization with M lineage bees seemed to have resulted in lowered population diversity from the perspective of the progenitor A lineage bees, it is clear that the A lineage potentially benefitted in this genetic exchange through the inheritance and maintenance

of specific M lineage genes. If the M lineage provided specific benefits to the population, these loci would be maintained across the population. Conversely, if no benefits were conferred at the time of hybridization, loci would be randomly distributed throughout the genome. Using local ancestry analysis, we identified 89 genomic intervals across the *A. m. intermissa* genome that were enriched for M lineage ancestry. By examining these segments, we gained a deeper understanding of the genomic contributions of the M lineage within this hybrid population. The average interval length was 28,200 bp, with the longest segment measuring 412,423 bp on chromosome 11. This is interesting because in hybrid A and M lineage populations, chromosome 11 has previously been identified as a region where M-lineage enrichment may confer adaptive advantages to hybrids (Calfee et al., 2020; Everitt et al., 2023; Nelson et al., 2017). In Brazilian African-European hybrid populations, a conserved M lineage enriched segment spanning 14.0–15.4 Mbp was identified, overlapping with QTLs related to foraging behavior, ovary size, and ovariole number (Nelson et al., 2017). Similarly, a study on Californian African-European hybrid populations identified a comparable region on chromosome 11 from 13.9–15.3 Mbp (Calfee et al., 2020). In contrast, the M lineage enriched regions identified in this study show no overlap with previously reported enriched M lineage regions in African-European hybrid populations. Notably, the largest M-lineage-enriched region in the *A. m. intermissa* population was identified on chromosome 11, spanning 6.84–7.24 Mbp, highlighting this chromosome as a potential hotspot for M-lineage enrichment in African–European hybrid populations. Interestingly, the largest number of regions enriched for M lineage contribution in the Moroccan *A. m. intermissa* population were located on chromosome 7, which comprised 25.84% of all enriched intervals. Chromosome 7 may play a significant role in the Moroccan *A. m. intermissa* population, warranting further investigation. Understanding the mechanisms underlying chromosome 11 and

chromosome 7 M lineage enrichment could provide valuable insights into hybrid adaptation in Northwest Africa.

Next, we sought to explore the functional significance of genomic M lineage enrichment in *A. m. intermissa*, to do this we identified the genes within enriched regions, then performed functional annotation analysis with DAVID v2024q4 (Huang et al., 2009; Sherman et al., 2022). The analysis found one cluster of significance. This cluster was found to be enriched for terms associated with juvenile hormone production, which is known to be a critical driver of division of labour in *A. mellifera* (Elekonich et al., 2001). The genetic enrichment of these functions raises the possibility that European genetic contributions may influence *A. m. intermissa's* life cycle timing, as colonies of this subspecies are known to follow a Mediterranean brood rhythm (Ruttner, 1988).

To further investigate the functional significance of M lineage gene enrichment in *A. m. intermissa*, we conducted a literature review to identify supporting studies and determine the biological processes associated with these genes. Within the Moroccan Hybrid *A. m. intermissa population*, M lineage genetic contributions were found to underly several genes with functions postulated to contribute to processes such as detoxification, immune response, development, and behavioral regulation.

Within detoxification pathways, We found that M lineage regions were enriched for the cytochrome P450 genes Cyp6be1 (*GB46814*) and Cyp6as3 (*GB49887*), both of which play key roles in xenobiotic detoxification (Berenbaum & Johnson, 2015; Johnson et al., 2018). These genes are known to be differentially expressed in response to neonicotinoid agrochemicals, including imidacloprid (Alptekin et al., 2016; Chen et al., 2021), suggesting a role in pesticide metabolism. Another detoxification-related gene within M lineage-enriched regions, Gtpx-2

(*GB48634*), encodes an enzyme involved in primary antioxidant response, acting directly on reactive oxygen species (ROS). This gene has been shown to be upregulated in response to nicotine exposure in honey bees (Corona & Robinson, 2006; Rand et al., 2015), supporting the role of M lineage segments in mitigating chemical stressors.

The M lineage contributions also intersect with genes involved in immune response and colony regulation. Insulin gene enhancer protein (GB45757) is upregulated in honey bee workers in response to Varroa mite infections and is a key component of the insulin-like growth factor signaling pathway (Zanni et al., 2017). This gene has been implicated in both immune function and behavioral transitions, particularly in regulating the shift from nursing to foraging (Ament et al., 2008; Zanni et al., 2017), suggesting that M lineage influence may extend to honey bee division of labor and resilience to parasites.

Several genes overlapping with M lineage segments were found to be associated with cuticle formation and structural development, critical for exoskeleton integrity and pigmentation. The gene GB45596 was found to support the fatty acid elongation process in honey bees and, in a transcriptomics study, was shown to peak in newly emerged adults and foragers. (Falcón et al., 2014; Falcon et al., 2019). Additionally, Grainy head (GB46725) an ecdysone-response transcription factor, was found within M lineage-enriched regions in this population. Grainy head plays a well-documented role in cuticle differentiation across insect species, including *Drosophila* (Gangishetti et al., 2012). Another well documented gene located within M lineage enriched segments, ebony (GB46429), is involved in pigmentation pathways, influencing the development of yellow coloration (True, 2003; Wittkopp & Beldade, 2009). The overlap of these genes with M lineage segments suggests that the M lineage contribution may influence various functions underlying structural and pigment formation in this Moroccan population.

M lineage contributions may extend to supporting aspects of honey bee biology associated with neurodevelopment and behavioral regulation, as evidenced by relevant studies on six major genes within these M lineage–enriched segments. *Abscam* (GB45774), a *Drosophila dscam2* ortholog, supports neuronal outgrowth and patterning, potentially influencing complex neural architecture required for navigation and social behavior (Funada et al., 2007). Meanwhile, *PAX6* (GB50342), the *Drosophila* ortholog of *eyeless*, is a transcription factor involved in regulating the transition from nursing to foraging, reinforcing its role in worker division of labor (Khamis et al., 2015). These genes suggest that M lineage segments contribute to cognitive and behavioral adaptations relevant to social organization in *A. m. intermissa*.

Further evidence of M lineage influence is found in genes involved in hormonal signaling and developmental regulation. *Ethr (*GB46500) encodes the ecdysis-triggering hormone receptor (ETHR), essential for molting and juvenile hormone regulation in insects (Areiza et al., 2014; Noriega, 2014). Another key regulator, FOXO transcription factor (GB48301), integrates environmental cues to modulate juvenile hormone pathways, influencing caste differentiation and physiological adaptation (Alvarado et al., 2015; Birkenkamp & Coffer, 2003; Kamakura, 2011; Wang et al., 2020). Additionally, GB55060 (*Rhomboid-related protein 3-like*) and GB55364 (*tyrosine-protein phosphatase*) have putative roles in signal transduction and cellular differentiation, linking M lineage contributions to developmental plasticity (Bier et al., 1990; McQuibban et al., 2006; Tonks, 2006).

Finally, M lineage segments were enriched for hexokinase2 (GB47079), a key enzyme in glycolysis, which has been linked to colony defense behaviors in honey bees. Genes involved in sugar metabolism are frequently associated with aggression and defensive responses, making this finding particularly notable (Bier et al., 1990; Harpur et al., 2020; Li-Byarlay et al., 2014).

Additionally, of the 89 loci enriched for the M lineage, 18 segments showed complete overlap within 4 QTLs, which were all determined to be associated with varroa sensitivity hygiene (Spötter et al., 2012). Hygienic behavior can be understood as the ability for *A. mellifera* to detect and remove diseased brood from the colony, and varroa sensitivity hygiene is a more specialized form of this trait specific to the detection and removal of brood infected with *V. destructor*. The enrichment of the M lineage within these QTLs suggests it may confer an adaptive advantage, benefiting the population's phenotype for varroa sensitivity hygiene. This research enhances our understanding of previously published work on the subspecies, which determined that *A. m. intermissa* demonstrates an increased resistance to *Varroa destructor*, attributed to its grooming behavior (Adjlane et al., 2016; Boecking & Ritter, 1993; Dadoun et al., 2020; Locke, 2016).

Conclusions

Through global ancestry and diversity analyses, combined with admixture time estimation, we confirmed that this Moroccan population of *A. m. intermissa* is a recently admixed population with predominant contributions from African (A) and European (M) lineages. Our analyses indicate that this admixture occurred 14.04 generations ago. Local ancestry analyses identified genomic regions enriched for M lineage, notably on chromosome 7, which contained the highest level of M lineage enrichment across all chromosomes, and chromosome 11, which harbored the longest uninterrupted M lineage genomic segment. The unique pattern of M lineage enrichment at these locations suggests they play an important role in conferring fitness benefits to hybrids, creating an avenue for exploration in future work. Enrichment analysis highlighted a significant association between M lineage gene enrichment and biological processes related to juvenile hormone production, a key regulator of division of

labor in *Apis mellifera*. Supporting analysis by gene-level exploration highlighted contributions to detoxification, immune responses, development, and colony behavior. Additionally, one of *A. m. intermissa's* most agriculturally valuable phenotypes, its resilience to *Varroa destructor*, may be influenced by 4 QTLS containing 18 loci contributed from the M lineage in this hybrid population. These findings underscore the diverse roles of European lineage genetics in African honey bee populations, which may enhance adaptability and resilience in hybrid populations. This work establishes a foundation for future studies to investigate the ecological and evolutionary significance of admixture in *A. m. intermissa* and its potential parallels in other hybrid honeybee populations.

Future Directions

To advance this work, future studies should broaden the representation of *A. m. intermissa* by sampling across its full native range, including Morocco, Algeria, and Tunisia (Haddad et al., 2015; Ruttner, 1988). This would enable more comprehensive population-level analyses and shed light on how admixture patterns vary geographically. Subspecies preservation remains a key focus in North African honey bee research, with studies examining admixture levels between *A. m. intermissa* and *A. m. sahariensis (Aglagane et al., 2023)*. Continued investigation of *Apis mellifera* population structure across North Africa, particularly the interaction between *A. m. intermissa* and *A. m. sahariensis*, will support ongoing conservation by monitoring introgression levels over time.

This work sought to provide a starting point for estimating the timeline of admixture events between African and European honey bees in Morocco. To do this, we assume a single-pulse admixture model (Moorjani & Hellenthal, 2023; Moorjani et al., 2011). To refine our understanding of the timing and nature of the event, complementary analyses could test whether the admixture reflects a single pulse, multiple pulses, or continuous gene flow. This would provide an increased resolution on date estimate of the admixture event that occurred within this Moroccan *Apis mellifera intermissa* population. (Moorjani & Hellenthal, 2023; Moorjani et al., 2011).

Furthermore, to build on the local ancestry analysis of the *A. m. intermissa* population, simulations of population admixture in the absence of selection, tailored to reflect the admixture model and fine-scale timing estimates previously inferred for this *Apis mellifera intermissa* population, could be used to test whether the observed patterns are consistent with neutral processes or suggest adaptive retention of ancestry segments (Nelson et al., 2017).

Additionally, unique admixture hotspots identified on chromosomes 7 and 11 should be investigated for potential functional significance. This could help determine whether these regions contribute to adaptive traits in this Admixed *A. m. intermissa* population. Future work might include transcriptomic profiling or selection scans (Saravanan et al., 2020; Tzec-Interián et al., 2025).

Overall, this work serves as a foundation for assessing the admixture patterns of the notably admixed *Apis mellifera intermissa* population. It offers valuable insights into the history and adaptive potential of this European–African hybrid honey bee, while also opening multiple avenues for further research into genomic admixture and the biogeography of honey bees.

Supplemental Information

Table S1. Characteristics of the *Apis mellifera* genomes utilized in this study, including sampling locations and NCBI Database identifiers.

Sample Name ca	Subspecies	Location	BioProject	BioSample	SRA Number
Scut_A_82	scutellata	South Africa: East Cape	PRJNA216922	SAMN02324189	SRR957075
Scut_A_83	scutellata	South Africa: East Cape	PRJNA216922	SAMN02324190	SRR957076
Scut_A_84	scutellata	South Africa: NW. Cape	PRJNA216922	SAMN02324191	SRR957077
Scut_A_85	scutellata	South Africa: N. Cape	PRJNA216922	SAMN02324192	SRR957078
Scut_A_86	scutellata	South Africa: NW. Cape	PRJNA216922	SAMN02324193	SRR957091
Scut_A_87	scutellata	South Africa: West Cape	PRJNA216922	SAMN02324194	SRR957092
Scut_A_88	scutellata	South Africa: Freestate	PRJNA216922	SAMN02324195;	SRR957093
Scut_A_89	scutellata	South Africa: N/NW Cape	PRJNA216922	SAMN02324196	SRR957094
Cauca_O_1	caucasica	Turkey: Ardahan	PRJNA729035	SAMN19109567	SRR14703849
Cauca_O_2	caucasica	Turkey: Ardahan	PRJNA729035	SAMN19109568	SRR14703848
Cauca_O_3	caucasica	Turkey: Posov	PRJNA729035	SAMN19109681	SRR14703847
Cauca_O_4	caucasica	Turkey: Posov	PRJNA729035	SAMN19109682	SRR14703846
Iberi_M_1	iberiensis	Spain: Cordoba	PRJNA729035	SAMN19109596	SRR14703845
Iberi_M_2	iberiensis	Spain: Cordoba	PRJNA729035	SAMN19109597	SRR14703844
Iberi_M_3	iberiensis	Spain: Cordoba	PRJNA729035	SAMN19109598	SRR14703843
Iberi_M_4	iberiensis	Spain: Lugo	PRJNA729035	SAMN19109632	SRR14703842
Scut_A_90	scutellata	South Africa: NW. Cape	PRJNA216922	SAMN02324197	SRR957095
Scut_A_91	scutellata	South Africa: East Cape	PRJNA216922	SAMN02324198	SRR957096
Scut_A_92	scutellata	South Africa: East Cape	PRJNA216922	SAMN02324199	SRR957097

Scut_A_56	scutellata	Kenya: Mt Kenya savannah lowland	PRJNA357367	SAMN06142946	SRR5115494
Scut_A_57	scutellata	Kenya: Mt Kenya savannah lowland	PRJNA357367	SAMN06142942	SRR5115508
Scut_A_58	scutellata	Kenya: Mt Kenya savannah lowland	PRJNA357367	SAMN06142943	SRR5115510
Scut_A_59	scutellata	Kenya: Mt Kenya savannah lowland	PRJNA357367	SAMN06142944	SRR5115502
Scut_A_60	scutellata	Kenya: Mt Kenya savannah lowland	PRJNA357367	SAMN06142947	SRR5115518
Scut_A_61	scutellata	Kenya: Mt Kenya savannah lowland	PRJNA357367	SAMN06142948	SRR5115504
Scut_A_62	scutellata	Kenya: Mt Kenya savannah lowland	PRJNA357367	SAMN06142949	SRR5115512
Scut_A_63	scutellata	Kenya: Mt Kenya savannah lowland	PRJNA357367	SAMN06142950	SRR5115509
Scut_A_64	scutellata	Kenya: Mau savannah lowland	PRJNA357367	SAMN06142928	SRR5115490
Scut_A_65	scutellata	Kenya: Mau savannah lowland	PRJNA357367	SAMN06142929	SRR5115506
Scut_A_66	scutellata	Kenya: Mau savannah lowland	PRJNA357367	SAMN06142930	SRR5115521
Scut_A_67	scutellata	Kenya: Mau savannah lowland	PRJNA357367	SAMN06142931	SRR5115501
Lamark_L_2	lamarckii	Egypt: Dairout	PRJNA729035	SAMN19109607	SRR14703822
Lamark_L_4	lamarckii	Egypt: Dairout	PRJNA729035	SAMN19109610	SRR14703816
Lamark_L_8	lamarckii	Egypt: Assiut	PRJNA729035	SAMN19109578	SRR14703810
Lamark_L_9	lamarckii	Egypt: Assiut	PRJNA729035	SAMN19109579	SRR14703809
Lamark_L_10	lamarckii	Egypt: Assiut	PRJNA729035	SAMN19109580	SRR14703808
Lamark_L_11	lamarckii	Egypt: Assiut	PRJNA729035	SAMN19109582	SRR14703805
Lamark_L_12	lamarckii	Egypt: Assiut	PRJNA729035	SAMN19109583	SRR14703804
Lamark_L_13	lamarckii	Egypt: Assiut	PRJNA729035	SAMN19109584	SRR14703803
Ligust_C_1	ligustica	Italy: Milano	PRJNA729035	SAMN19109661	SRR14703802
Ligust_C_2	ligustica	Italy: Milano	PRJNA729035	SAMN19109662	SRR14703801
Ligust_C_3	ligustica	Italy: Milano	PRJNA729035	SAMN19109663	SRR14703800

Ligust_C_4	ligustica	Italy: Milano	PRJNA729035	SAMN19109664	SRR14703799
Ligust_C_5	ligustica	Italy: Bolongna	PRJNA729035	SAMN19109665	SRR14703798
Ligust_C_6	ligustica	Italy: Bolongna	PRJNA729035	SAMN19109666	SRR14703797
Melif_M_1	mellifera	France: Ille Sur Tet	PRJNA729035	SAMN19109629	SRR14703795
Melif_M_2	mellifera	France: Ille Sur Tet	PRJNA729035	SAMN19109630	SRR14703794
Melif_M_3	mellifera	France: Avignon	PRJNA729035	SAMN19109585	SRR14703793
Melif_M_4	mellifera	France: Avignon	PRJNA729035	SAMN19109586	SRR14703792
Melif_M_5	mellifera	France: Landes	PRJNA729035	SAMN19109631	SRR14703791
Anato_O_6	anatoliaca	Turkey: Menemen	PRJNA729035	SAMN19109722	SRR14703783
Anato_O_7	anatoliaca	Turkey: Menemen	PRJNA729035	SAMN19109723	SRR14703782
Anato_O_8	anatoliaca	Turkey: Fethiye	PRJNA729035	SAMN19109724	SRR14703781
Anato_O_9	anatoliaca	Turkey: Fethiye	PRJNA729035	SAMN19109718	SRR14703780
Anato_O_10	anatoliaca	Turkey: Fethiye	PRJNA729035	SAMN19109719	SRR14703779
Anato_O_11	anatoliaca	Turkey: Fethiye	PRJNA729035	SAMN19109720	SRR14703778
Anato_O_12	anatoliaca	Turkey: Fethiye	PRJNA729035	SAMN19109721	SRR14703777
Scut_A_68	scutellata	Kenya: Mau savannah lowland	PRJNA357367	SAMN06142922	SRR5115520
Scut_A_69	scutellata	Kenya: Mau savannah lowland	PRJNA357367	SAMN06142923	SRR5115489
Scut_A_70	scutellata	Kenya: Mau savannah lowland	PRJNA357367	SAMN06142924	SRR5115522
Scut_A_71	scutellata	Kenya: Mau savannah lowland	PRJNA357367	SAMN06142925	SRR5115515
Scut_A_72	scutellata	Kenya: Mau savannah lowland	PRJNA357367	SAMN06142926	SRR5115527
Scut_A_73	scutellata	Kenya: Mau savannah lowland	PRJNA357367	SAMN06142927	SRR5115516
A_Adan_1	adansonii	Burkina Faso: Kurkina	PRJNA729035	SAMN19109587	SRR14703898
A_Adan_2	adansonii	Burkina Faso: Kurkina	PRJNA729035	SAMN19109588	SRR14703897
A_Adan_3	adansonii	Burkina Faso: Kurkina	PRJNA729035	SAMN19109589	SRR14703851
A_Adan_4	adansonii	Burkina Faso: Kurkina	PRJNA729035	SAMN19109590	SRR14703840

A_Adan_5	adansonii	Burkina Faso: Kurkina	PRJNA729035	SAMN19109591	SRR14703829
A_Adan_6	adansonii	Burkina Faso: Kurkina	PRJNA729035	SAMN19109592	SRR14703818
A_Adan_7	adansonii	Burkina Faso: Kurkina	PRJNA729035	SAMN19109593	SRR14703807
A_Adan_8	adansonii	Senegal	PRJNA729035	SAMN19109689	SRR14703796
Inter_A_7	Intermissa	Morocco: Asilah	PRJNA729035	SAMN19109569	SRR14703841
Inter_A_8	Intermissa	Morocco: Asilah	PRJNA729035	SAMN19109570;	SRR14703839
Inter_A_9	Intermissa	Morocco: Asilah	PRJNA729035	SAMN19109571	SRR14703838
Inter_A_10	Intermissa	Morocco: Asilah	PRJNA729035	SAMN19109572	SRR14703837
Inter_A_11	Intermissa	Morocco: Asilah	PRJNA729035	SAMN19109573	SRR14703836
Inter_A_12	Intermissa	Morocco: Asilah	PRJNA729035	SAMN19109574	SRR14703835
Inter_A_13	Intermissa	Morocco: Asilah	PRJNA729035	SAMN19109575	SRR14703834
Inter_A_14	Intermissa	Morocco: Asilah	PRJNA729035	SAMN19109576	SRR14703833
Inter_A_15	Intermissa	Morocco	PRJNA729035	SAMN19109667	SRR14703832
Jementi_Y_37	jemenitica	Yemen: Sanaa	PRJNA729035	SAMN19109687	SRR14703858
Jementi_Y_38	jemenitica	Yemen: Sanaa	PRJNA729035	SAMN19109688	SRR14703857
carni_C_41	carnica	Croatia: Novalija, Pag Island	PRJNA216922	SAMN02324163	SRR957082
carni_C_42	carnica	Croatia: Novalija, Pag Island	PRJNA216922	SAMN02324164	SRR957083
carni_C_43	carnica	Croatia: Baric Draga	PRJNA216922	SAMN02324165	SRR957084
carni_C_44	carnica	Croatia: Sugarje	PRJNA216922	SAMN02324166	SRR957085
carni_C_45	carnica	Slovenia	PRJNA216922	SAMN02324167	SRR957086
carni_C_46	carnica	Slovenia	PRJNA216922	SAMN02324169	SRR957088
Melif_M_47	mellifera	Poland: Zalewo	PRJNA216922	SAMN02324170	SRR957089
Melif_M_48	mellifera	Poland: Rudawka	PRJNA216922	SAMN02324171	SRR957090
Melif_M_49	mellifera	Poland: Rudawka	PRJNA216922	SAMN02324172	SRR957058
Melif_M_50	mellifera	Poland: Sucha Rzeczka	PRJNA216922	SAMN02324173	SRR957059
Melif_M_51	mellifera	Poland: Sucha Rzeczka	PRJNA216922	SAMN02324174	SRR957060
Iberi_M_52	iberiensis	Spain: Cordoba	PRJNA216922	SAMN02324175	SRR957061
Iberi_M_53	iberiensis	Spain: Murcia	PRJNA216922	SAMN02324176	SRR957062

Iberi_M_54	iberiensis	Spain: Murcia	PRJNA216922	SAMN02324177	SRR957063
Iberi_M_55	iberiensis	Spain: Cordoba	PRJNA216922	SAMN02324178	SRR957064
Inter_A_16	Intermissa	Morocco	PRJNA729035	SAMN19109668	SRR14703831
Inter_A_17	Intermissa	Morocco	PRJNA729035	SAMN19109669	SRR14703830
Inter_A_18	Intermissa	Morocco	PRJNA729035	SAMN19109670	SRR14703828
Inter_A_19	Intermissa	Morocco	PRJNA729035	SAMN19109671	SRR14703827
Inter_A_20	Intermissa	Morocco	PRJNA729035	SAMN19109672	SRR14703826
Inter_A_21	Intermissa	Morocco	PRJNA729035	SAMN19109673	SRR14703825
Scut_A_1	scutellata	Kenya: Marchorwe	PRJNA729035	SAMN19109658	SRR14703927
Scut_A_2	scutellata	Kenya: Moorland at Mt. Elgon	PRJNA729035	SAMN19109678	SRR14703926
Scut_A_3	scutellata	Kenya: Moorland at Mt. Elgon	PRJNA729035	SAMN19109679	SRR14703925
Scut_A_4	scutellata	Kenya: Marchorwe	PRJNA729035	SAMN19109659	SRR14703924
Scut_A_5	scutellata	Kenya: Malewa	PRJNA729035	SAMN19109641	SRR14703923
Scut_A_6	scutellata	Kenya: Malewa	PRJNA729035	SAMN19109642	SRR14703922
Scut_A_7	scutellata	Kenya: Malewa	PRJNA729035	SAMN19109643	SRR14703921
Scut_A_8	scutellata	Kenya: Gede Ruins	PRJNA729035	SAMN19109617	SRR14703920
Scut_A_9	scutellata	Kenya: Mandera	PRJNA729035	SAMN19109650	SRR14703919
Scut_A_10	scutellata	Kenya: Mandera	PRJNA729035	SAMN19109651	SRR14703918
Scut_A_11	scutellata	Kenya: Mandera	PRJNA729035	SAMN19109652	SRR14703917
Scut_A_12	scutellata	Kenya: Marchorwe	PRJNA729035	SAMN19109660	SRR14703916
Jementi_Y_74	jemenitica	Saudi Arabia	PRJNA216922	SAMN02324180	SRR957066
Jementi_Y_75	jemenitica	Saudi Arabia	PRJNA216922	SAMN02324182	SRR957068
Jementi_Y_76	jemenitica	Saudi Arabia	PRJNA216922	SAMN02324183	SRR957069
Jementi_Y_77	jemenitica	Saudi Arabia	PRJNA216922	SAMN02324184	SRR957070
Jementi_Y_78	jemenitica	Saudi Arabia	PRJNA216922	SAMN02324185	SRR957071
Jementi_Y_79	jemenitica	Yemen	PRJNA216922	SAMN02324186	SRR957072
Jementi_Y_80	jemenitica	Yemen	PRJNA216922	SAMN02324187	SRR957073
Jementi_Y_81	jemenitica	Yemen	PRJNA216922	SAMN02324188	SRR957074
Scut_A_13	scutellata	Kenya: Ichiara	PRJNA729035	SAMN19109628	SRR14703914
Scut_A_14	scutellata	Kenya: Gede Ruins	PRJNA729035	SAMN19109618	SRR14703912
Scut_A_15	scutellata	Kenya: Tanzania border	PRJNA729035	SAMN19109711	SRR14703911

Scut_A_16	scutellata	Kenya: Tanzania border	PRJNA729035	SAMN19109712	SRR14703910
Scut_A_17	scutellata	Kenya: Tanzania border	PRJNA729035	SAMN19109713	SRR14703909
Scut_A_18	scutellata	Botswana: Gaborone	PRJNA729035	SAMN19109616	SRR14703908
Scut_A_19	scutellata	Kenya: Mandera	PRJNA729035	SAMN19109647	SRR14703907
Scut_A_20	scutellata	Kenya: Mandera	PRJNA729035	SAMN19109648	SRR14703906
Scut_A_21	scutellata	Kenya: Mandera	PRJNA729035	SAMN19109649	SRR14703905
Scut_A_22	scutellata	Zimbabwe: Harare	PRJNA729035	SAMN19109622	SRR14703904
Scut_A_23	scutellata	Zimbabwe: Harare	PRJNA729035	SAMN19109624	SRR14703901

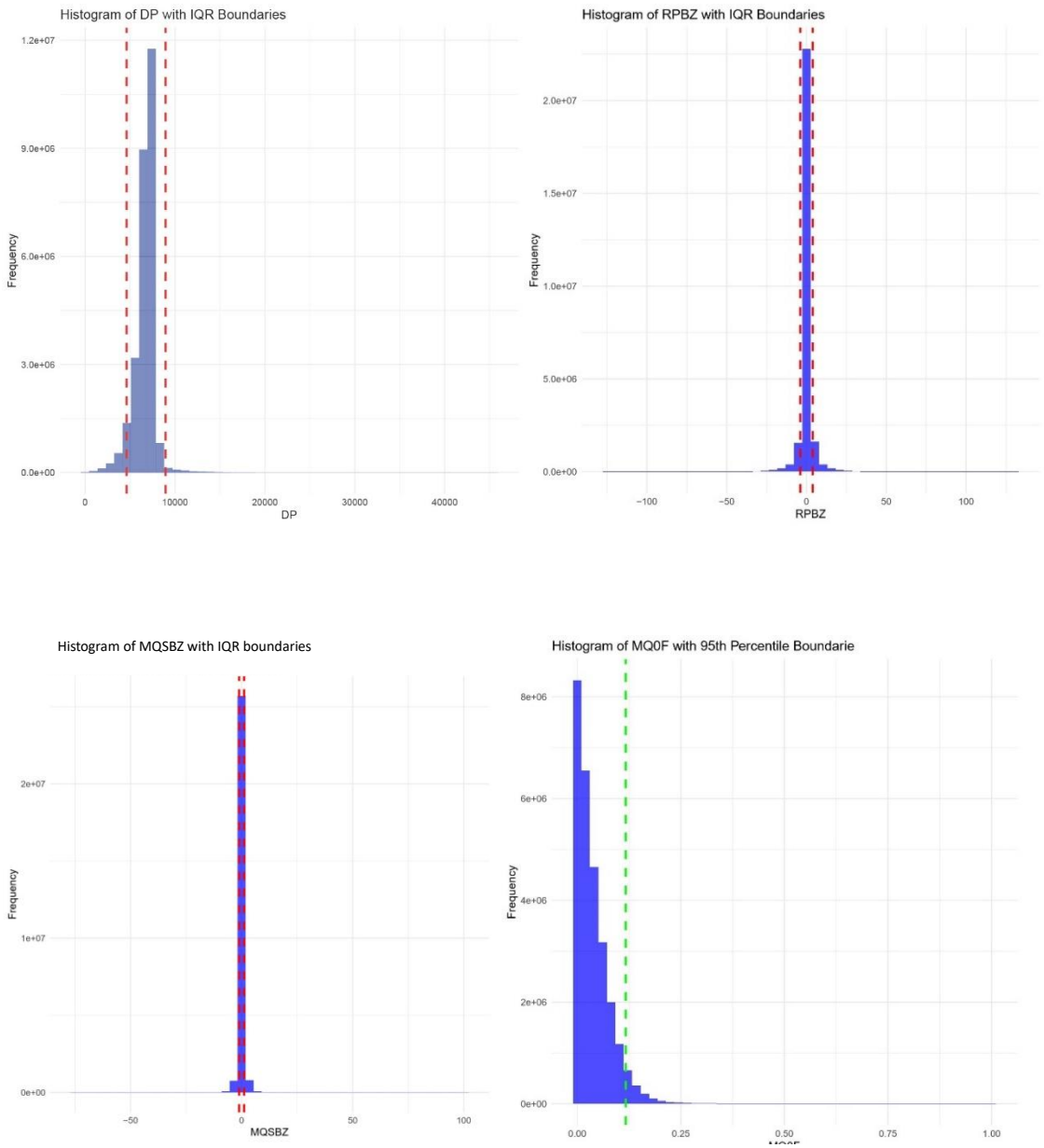

Figure S1. Distributions of SNP Metrics with Filter Thresholds. This table presents the distributions of key SNP quality metrics including depth (DP), MQSBZ (Mapping Quality Score Bias Z-score), RPBZ (Read Position Bias Z-score), and MQ0F (Mapping Quality Zero Fraction). The lines in the table represent the specific filter thresholds applied: depth between 5,642 and 11,322, MQSBZ from -1.204642 to 1.216214, RPBZ from -3.998202 to 4.047033, and MQ0F at 0.0921153 or lower. These thresholds were set to exclude low-quality SNPs from further analysis, ensuring the retention of high-quality genomic data.

Figure S2. Ancestry proportions of *Apis mellifera* lineages inferred through global admixture analysis of N = 134 samples across a range of K values (K = 6 to K = 8). Each vertical bar represents an individual bee, with colors indicating the proportion of its genome attributed to each inferred ancestral population. Increasing K values reveal finer-scale population structure, including the emergence of an A. m. intermissa-specific cluster at K = 6–8 and the separation of *A. m. lamarkii* from the broader African lineage at K = 7.

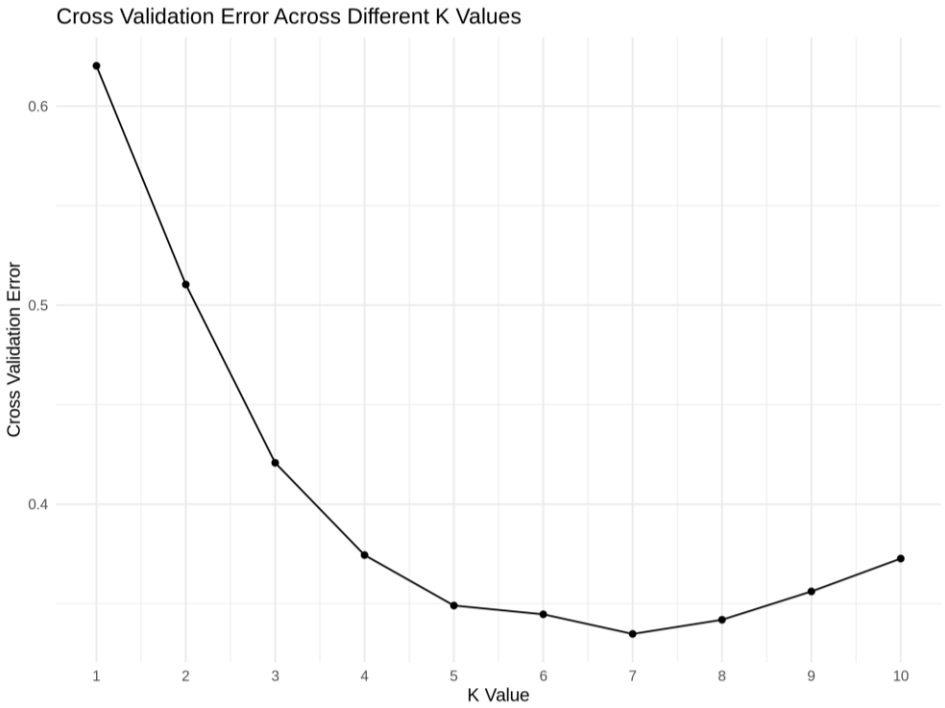

Figure S3. *Visualization of K-values (1–10) based on 10-fold cross-validation error for ADMIXTURE analysis of 134 Apis mellifera samples. The plot illustrates the cross-validation error for each K-value, with the lowest error indicating the optimal number of ancestral populations for clustering genetic variation.*

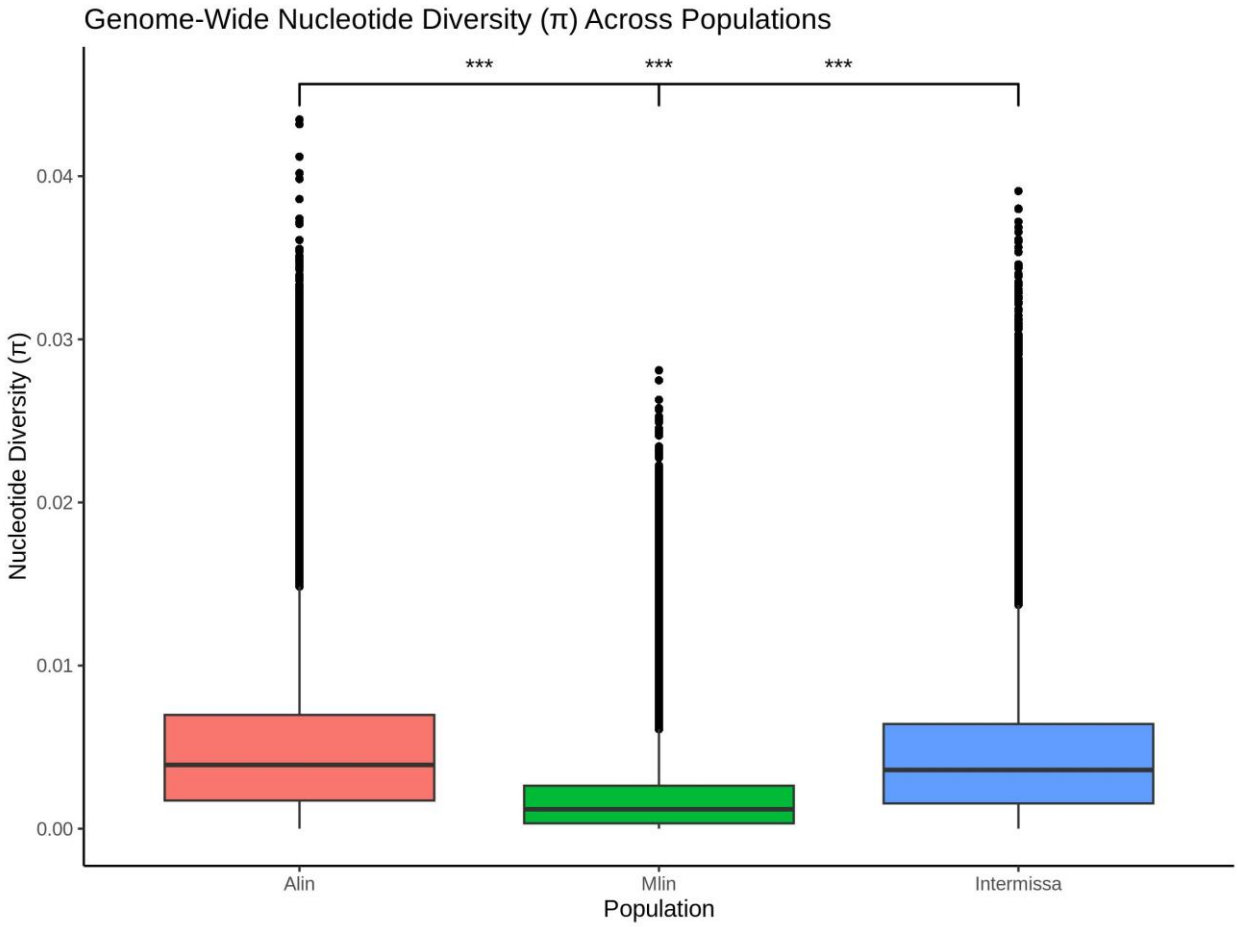

Figure S4. Box plot of nucleotide diversity (π) values across the A, M, and admixed A. m. intermissa populations. Lineage A exhibits the highest genome-wide nucleotide diversity (π = 0.0039), while lineage M shows the lowest (π = 0.0012), with A. m. intermissa displaying intermediate values (π = 0.0036). Statistical analysis using A Friedman test confirmed significant differences among the three groups (χ^2 = 752,328, df = 2, $p < 0.001$) with post-hoc Wilcoxon signed-rank tests revealing significant pairwise differences ($p < 0.0001$).

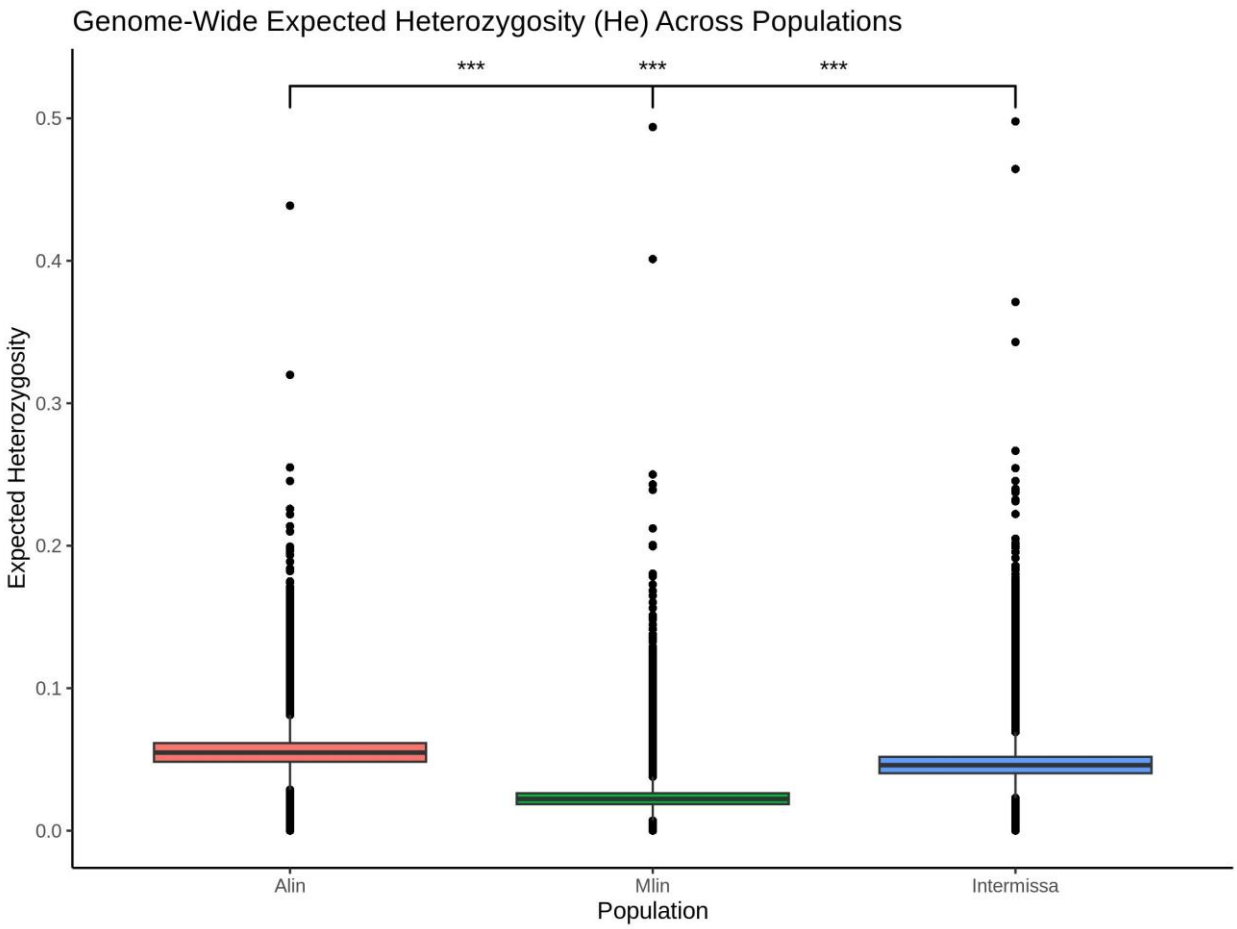

Figure S5. Box plot of expected heterozygosity values for the A, M, and admixed A. m. intermissa populations. Expected heterozygosity was highest in lineage A (H_e = 0.0548), followed by A. m. intermissa (H_e = 0.0460), with the lowest values recorded in lineage M (H_e = 0.0223). Friedman test results showed highly significant differences among the groups (χ^2 = 1,454,544, df = 2, $p < 0.001$) and post-hoc Wilcoxon signed-rank tests confirmed significant pairwise differences ($p < 0.001$).

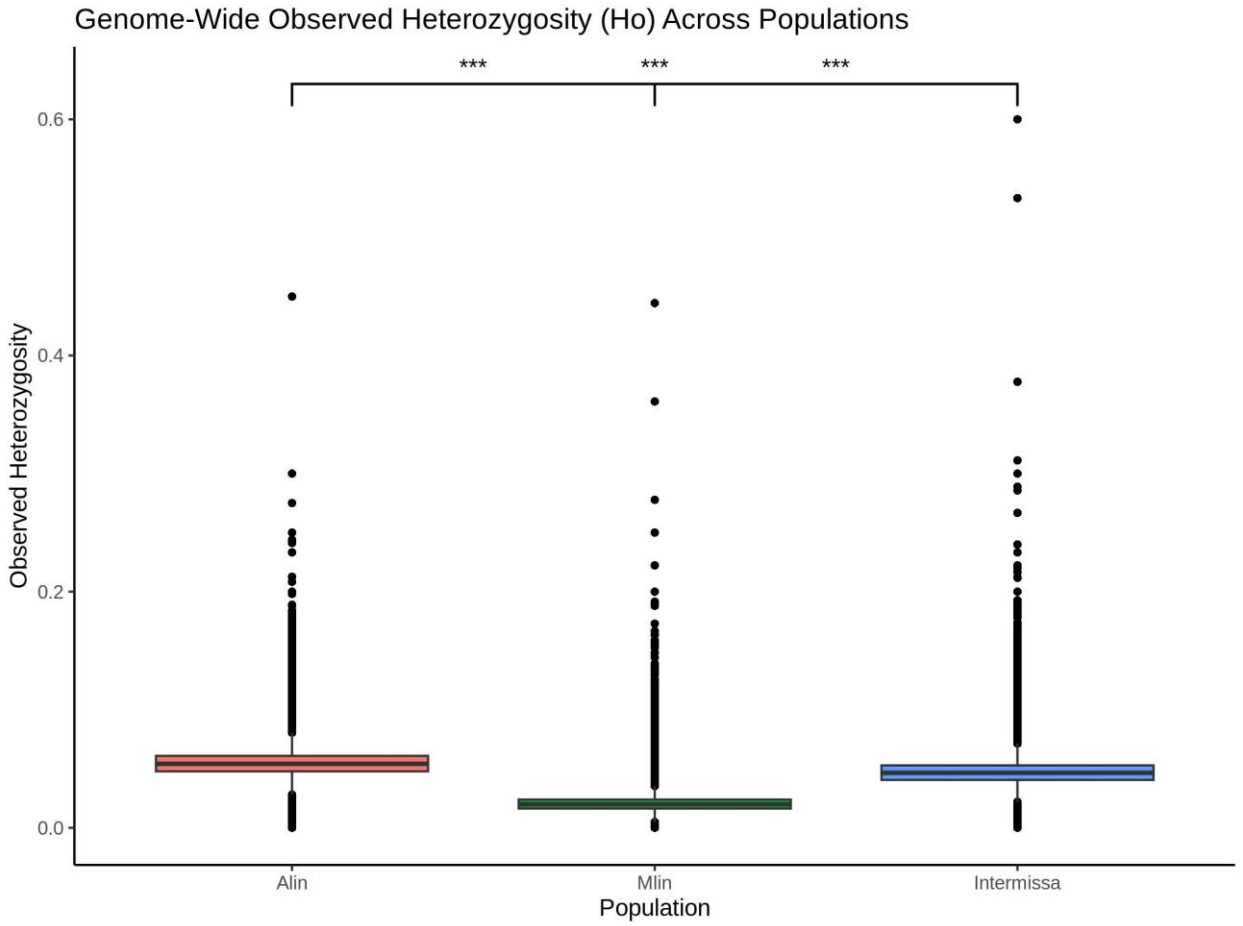

Figure S6. Box plot of observed heterozygosity values for the A, M, and admixed *A. m. intermissa* populations. Observed heterozygosity followed the same trend as expected heterozygosity, with the highest values in lineage A (H_o = 0.0542), intermediate values in *A. m. intermissa* (H_o = 0.0465), and the lowest in lineage M (H_o = 0.0119). Friedman test confirmed significant differences among the groups (V = 347,551,619,906.0, p < 0.001). with post-hoc Wilcoxon signed-rank tests indicating significant pairwise differences (p < 0.001).

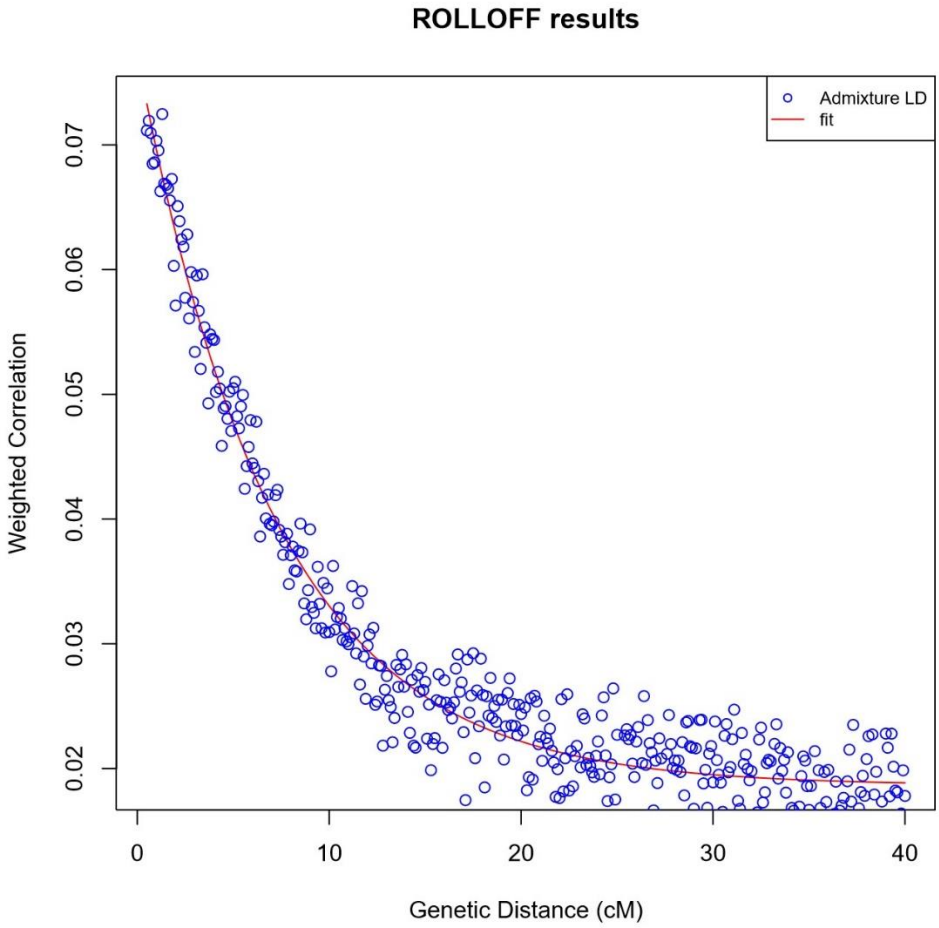

Figure S7. Estimation of Generations Since Admixture Using ROLLOFF. This figure illustrates the analysis performed using the ROLLOFF method to estimate the number of generations since the admixture event in Apis mellifera intermissa. The variant set enriched for ancestry-informative markers (AIMs) was analyzed to determine linkage disequilibrium (LD) decay over time within the admixed population. Ancestry decay was fitted to an exponential distribution using least-squares regression. This analysis estimated 14.04 generations have occurred since A and M population admixture.

Summary of fit:

Formula: wcorr ~ (C + A * exp(-m * dist/100))

Parameters:

	Estimate	Std. Error	t value	Pr(>|t|)	
C	1.864e-02	2.382e-04	78.24	<2e-16	***
A	5.864e-02	7.173e-04	81.74	<2e-16	***
m	1.404e+01	3.253e-01	43.17	<2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.002662 on 393 degrees of freedom

Number of iterations to convergence: 1

Achieved convergence tolerance: 1.061e-06

Estimated number of generations since mixture: 14.04

Supplemental S7 Continued Estimations of Generations Since Admixture Using ROLLOFF: the following information is the model fit output for the estimation of generations since admixture between the A and M population calculated using Rolloff.

Table S2. *Functional Enrichment Analysis of Genes Enriched for M Ancestry. This table summarizes the DAVID v2024q4 annotation analysis of genomic regions enriched for M ancestry. The Annotation Cluster (Enrichment Score: 2.23) is linked to pheromone synthesis.*

Annotation Cluster 1 Enrichment Score: 2.23				
Category	**Description**	**Count**	**P_Value**	**Benjamini**
GOTERM_BP_DIRECT	farnesyl diphosphate biosynthetic process	3	2.30E-03	4.60E-01
INTERPRO	FPS1-like	3	2.30E-03	5.70E-01
GOTERM_MF_DIRECT	dimethylallyltranstransferase activity	3	2.50E-03	1.20E-01
GOTERM_MF_DIRECT	geranyltranstransferase activity	3	2.50E-03	1.20E-01
GOTERM_BP_DIRECT	pheromone biosynthetic process	3	7.80E-03	7.70E-01
INTERPRO	Polyprenyl_synt	3	7.80E-03	6.20E-01
INTERPRO	Isoprenoid_synthase_dom_sf	3	1.00E-02	6.40E-01
KEGG_PATHWAY	Terpenoid backbone biosynthesis	3	6.80E-02	1.00E+00

.

References

Abou-Shaara, H. F., Al-Ghamdi, A. A., Khan, K. A., & Al-Kahtani, S. N. (2021). Genetic network analysis between Apis mellifera subspecies based on mtDNA argues the purity of specimens from North Africa, the Levant and Saudi Arabia. *Saudi Journal of Biological Sciences*, *28*(5), 2718-2725.

Adavoudi, R., & Pilot, M. (2021). Consequences of hybridization in mammals: A systematic review. *Genes*, *13*(1), 50.

Adjlane, N., Dainat, B., Gauthier, L., & Dietemann, V. (2016). Atypical viral and parasitic pattern in Algerian honey bee subspecies Apis mellifera intermissa and A. m. sahariensis. *Apidologie*, *47*, 631-641.

Aglagane, A., Oleksa, A., Er-Rguibi, O., Tofilski, A., El Mouden, E. H., Aamiri, A., & Aourir, M. (2023). Genetic diversity and population structure of the Saharan honey bee Apis mellifera sahariensis from southeastern Morocco: introgression assessment and implications for conservation. *Apidologie*, *54*(3), 31.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655-1664.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*, *19*(9), 1655-1664. https://doi.org/10.1101/gr.094052.109

Alptekin, S., Bass, C., Nicholls, C., Paine, M. J., Clark, S. J., Field, L., & Moores, G. D. (2016). Induced thiacloprid insensitivity in honeybees (A pis mellifera L.) is associated with up-regulation of detoxification genes. *Insect Molecular Biology*, *25*(2), 171-180.

Alvarado, S., Rajakumar, R., Abouheif, E., & Szyf, M. (2015). Epigenetic variation in the Egfr gene generates quantitative variation in a complex trait in ants. *Nature communications*, *6*(1), 6513.

Ament, S. A., Corona, M., Pollock, H. S., & Robinson, G. E. (2008). Insulin signaling is involved in the regulation of worker division of labor in honey bee colonies. *Proceedings of the National Academy of Sciences*, *105*(11), 4226-4231.

Areiza, M., Nouzova, M., Rivera-Perez, C., & Noriega, F. G. (2014). Ecdysis triggering hormone ensures proper timing of juvenile hormone biosynthesis in pharate adult mosquitoes. *Insect biochemistry and molecular biology*, *54*, 98-105.

Armstrong, E. E., Mooney, J. A., Solari, K. A., Kim, B. Y., Barsh, G. S., Grant, V. B., Greenbaum, G., Kaelin, C. B., Panchenko, K., & Pickrell, J. K. (2024). Unraveling the genomic diversity and admixture history of captive tigers in the United States. *Proceedings of the National Academy of Sciences*, *121*(39), e2402924121.

Batra, S. (1995). Bees and pollination in our changing environment. *Apidologie*, *26*(5), 361-370.

Bayer Global. (2023). *The Value of Pollinators*

To the Ecosystem and Our Economy. https://www.bayer.com/en/agriculture/article/economic-value-pollinators

Berenbaum, M. R., & Johnson, R. M. (2015). Xenobiotic detoxification pathways in honey bees. *Current opinion in insect science*, *10*, 51-58.

Bier, E., Jan, L. Y., & Jan, Y. N. (1990). rhomboid, a gene required for dorsoventral axis establishment and peripheral nervous system development in Drosophila melanogaster. *Genes & development*, *4*(2), 190-203.

Birkenkamp, K., & Coffer, P. (2003). Regulation of cell survival and proliferation by the FOXO (Forkhead box, class O) subfamily of Forkhead transcription factors. *Biochemical Society Transactions*, *31*(1), 292-297.

Bloch, G., Francoy, T. M., Wachtel, I., Panitz-Cohen, N., Fuchs, S., & Mazar, A. (2010). Industrial apiculture in the Jordan valley during Biblical times with Anatolian honeybees. *Proceedings of the National Academy of Sciences*, *107*(25), 11240-11244.

Boecking, O., & Ritter, W. (1993). Grooming and removal behaviour of Apis mellifera intermissa in Tunisia against Varroa jacobsoni. *Journal of Apicultural Research*, *32*(3-4), 127-134.

Bowen, D. Q. (2009). Last Glacial Maximum. In V. Gornitz (Ed.), *Encyclopedia of Paleoclimatology and Ancient Environments* (pp. 493-495). Springer Netherlands. https://doi.org/10.1007/978-1-4020-4411-3_122

Browning, B. L., Tian, X., Zhou, Y., & Browning, S. R. (2021). Fast two-stage phasing of large-scale sequence data. *The American Journal of Human Genetics*, *108*(10), 1880-1890.

Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *The American Journal of Human Genetics*, *103*(3), 338-348.

Buerkle, C. A., & Lexer, C. (2008). Admixture as the basis for genetic mapping. *Trends in ecology & evolution*, *23*(12), 686-694. https://doi.org/10.1016/j.tree.2008.07.008

Calfee, E., Agra, M. N., Palacio, M. A., Ramírez, S. R., & Coop, G. (2020). Selection and hybridization shaped the rapid spread of African honey bee ancestry in the Americas. *PLOS Genetics*, *16*(10), e1009038.

Cánovas, F., De la Rúa, P., Serrano, J., & Galián, J. (2008). Geographical patterns of mitochondrial DNA variation in Apis mellifera iberiensis (Hymenoptera: Apidae). *Journal of Zoological Systematics and Evolutionary Research*, *46*(1), 24-30.

Carpenter, M. H., & Harpur, B. A. (2021). Genetic past, present, and future of the honey bee (Apis mellifera) in the United States of America. *Apidologie*, *52*, 63-79.

Chahbar, N., Munoz, I., Dall'Olio, R., De la Rúa, P., Serrano, J., & Doumandji, S. (2013). Population structure of North African honey bees is influenced by both biological and anthropogenic factors. *Journal of Insect Conservation*, *17*, 385-392.

Chávez-Galarza, J., Henriques, D., Johnston, J. S., Carneiro, M., Rufino, J., Patton, J. C., & Pinto, M. A. (2015). Revisiting the I berian honey bee (A pis mellifera iberiensis) contact zone: maternal and genome-wide nuclear variations provide support for secondary contact from historical refugia. *Molecular Ecology*, *24*(12), 2973-2992.

Chen, Y.-R., Tzeng, D. T., & Yang, E.-C. (2021). Chronic effects of imidacloprid on honey bee worker development—molecular pathway perspectives. *International journal of molecular sciences*, *22*(21), 11835.

Corona, M., & Robinson, G. (2006). Genes of the antioxidant system of the honey bee: annotation and phylogeny. *Insect Molecular Biology*, *15*(5), 687-701.

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

Cridland, J. M., Tsutsui, N. D., & Ramírez, S. R. (2017). The complex demographic history and evolutionary origin of the western honey bee, Apis mellifera. *Genome Biology and Evolution*, *9*(2), 457-472.

Crozier, R., & Page, R. (1985). On being the right size: male contributions and multiple mating in social Hymenoptera. *Behavioral Ecology and Sociobiology*, *18*, 105-115.

Crozier, R. H., & Fjerdingstad, E. J. (2001). Polyandry in social Hymenoptera—disunity in diversity? Annales Zoologici Fennici,

Dadoun, N., Nait-Mouloud, M., Mohammedi, A., & Sadeddine Zennouche, O. (2020). Differences in grooming behavior between susceptible and resistant honey bee colonies after 13 years of natural selection. *Apidologie*, *51*, 793-801.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., & Sherry, S. T. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2). https://doi.org/10.1093/gigascience/giab008

Dias-Alves, T., Mairal, J., & Blum, M. G. (2018). Loter: a software package to infer local ancestry for a wide range of species. *Molecular biology and evolution*, *35*(9), 2318-2326.

Dogantzis, K. A., Tiwari, T., Conflitti, I. M., Dey, A., Patch, H. M., Muli, E. M., Garnery, L., Whitfield, C. W., Stolle, E., & Alqarni, A. S. (2021). Thrice out of Asia and the adaptive radiation of the western honey bee. *Science Advances*, *7*(49), eabj2151.

Dogantzis, K. A., & Zayed, A. (2019). Recent advances in population and quantitative genomics of honey bees. *Current opinion in insect science*, *31*, 93-98.

Elekonich, M. M., Schulz, D. J., Bloch, G., & Robinson, G. E. (2001). Juvenile hormone levels in honey bee (Apis mellifera L.) foragers: foraging experience and diurnal variation. *Journal of Insect Physiology*, *47*(10), 1119-1125.

Everitt, T., Wallberg, A., Christmas, M. J., Olsson, A., Hoffmann, W., Neumann, P., & Webster, M. T. (2023). The genomic basis of adaptation to high elevations in Africanized honey bees. *Genome Biology and Evolution*, *15*(9), evad157.

Eynard, S. E., Windig, J. J., Leroy, G., van Binsbergen, R., & Calus, M. P. (2015). The effect of rare alleles on estimated genomic relationships from whole genome sequence data. *BMC genetics*, *16*, 1-12.

Falcón, T., Ferreira-Caliman, M. J., Nunes, F. M. F., Tanaka, É. D., do Nascimento, F. S., & Bitondi, M. M. G. (2014). Exoskeleton formation in Apis mellifera: Cuticular hydrocarbons profiles and expression of desaturase and elongase genes during pupal and adult development. *Insect biochemistry and molecular biology*, *50*, 68-81.

Falcon, T., Pinheiro, D. G., Ferreira-Caliman, M. J., Turatti, I. C., Abreu, F. C. P. d., Galaschi-Teixeira, J. S., Martins, J. R., Elias-Neto, M., Soares, M. P., & Laure, M. B. (2019). Exploring integument transcriptomes, cuticle ultrastructure, and cuticular hydrocarbons profiles in eusocial and solitary bee species displaying heterochronic adult cuticle maturation. *PLOS ONE*, *14*(3), e0213796.

Fortune Business Insights. (2022). *Market value of honey worldwide from 2019 to 2028*. https://www.fortunebusinessinsights.com/market-value-of-honey-worldwide-2019-to-2028

Frankham, R. (2002). *Introduction to conservation genetics*. Cambridge University Press.

Funada, M., Hara, H., Sasagawa, H., Kitagawa, Y., & Kadowaki, T. (2007). A honey bee Dscam family member, AbsCAM, is a brain-specific cell adhesion molecule with the neurite outgrowth activity which influences neuronal wiring during development. *European Journal of Neuroscience*, *25*(1), 168-180.

Gangishetti, U., Veerkamp, J., Bezdan, D., Schwarz, H., Lohmann, I., & Moussian, B. (2012). The transcription factor Grainy head and the steroid hormone ecdysone cooperate during differentiation of the skin of Drosophila melanogaster. *Insect Molecular Biology*, *21*(3), 283-295.

Haddad, N. J., Loucif-Ayad, W., Adjlane, N., Saini, D., Manchiganti, R., Krishnamurthy, V., AlShagoor, B., Batainh, A. M., & Mugasimangalam, R. (2015). Draft genome sequence of the Algerian bee Apis mellifera intermissa. *Genomics Data*, *4*, 24-25.

Han, F., Wallberg, A., & Webster, M. T. (2012a). From where did the W estern honeybee (A pis mellifera) originate? *Ecology and Evolution*, *2*(8), 1949-1957.

Han, F., Wallberg, A., & Webster, M. T. (2012b). From where did the Western honeybee (Apis mellifera)originate? *Ecology and Evolution*, *2*(8), 1949-1957. https://doi.org/10.1002/ece3.312

Harpur, B. A., Kadri, S. M., Orsi, R. O., Whitfield, C. W., & Zayed, A. (2020). Defense Response in Brazilian Honey Bees (Apis mellifera scutellata × spp.) Is Underpinned by Complex Patterns of Admixture. *Genome Biology and Evolution*, *12*(8), 1367-1377. https://doi.org/10.1093/gbe/evaa128

Harpur, B. A., Minaei, S., Kent, C. F., & Zayed, A. (2012). Management increases genetic diversity of honey bees via admixture. *Molecular Ecology*, *21*(18), 4414-4421. https://doi.org/10.1111/j.1365-294x.2012.05614.x

Hepburn, H. R., & Radloff, S. E. (2013). Honeybees of Africa.
Hewitt, G. M. (1988). Hybrid zones-natural laboratories for evolutionary studies. *Trends in ecology & evolution*, *3*(7), 158-167.
Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*, *4*(1), 44-57.
Johnson, R. M., Harpur, B. A., Dogantzis, K. A., Zayed, A., & Berenbaum, M. R. (2018). Genomic footprint of evolution of eusociality in bees: floral food use and CYPome "blooms". *Insectes Sociaux*, *65*, 445-454.
Jones, J. C., Wallberg, A., Christmas, M. J., Kapheim, K. M., & Webster, M. T. (2019). Extreme differences in recombination rate between the genomes of a solitary and a social bee. *Molecular biology and evolution*, *36*(10), 2277-2291.
Kamakura, M. (2011). Royalactin induces queen differentiation in honeybees. *Nature*, *473*(7348), 478-483.
Keller, S., & Taylor, D. (2010). Genomic admixture increases fitness during a biological invasion. *Journal of Evolutionary Biology*, *23*(8), 1720-1731.
Kent, R. B. (1988). The introduction and diffusion of the African honeybee in South America. *Yearbook of the Association of Pacific Coast Geographers*, *50*(1), 21-43.
Khamis, A. M., Hamilton, A. R., Medvedeva, Y. A., Alam, T., Alam, I., Essack, M., Umylny, B., Jankovic, B. R., Naeger, N. L., & Suzuki, M. (2015). Insights into the transcriptional architecture of behavioral plasticity in the honey bee Apis mellifera. *Scientific Reports*, *5*(1), 11136.
Kolbe, J. J., Glor, R. E., Rodríguez Schettino, L., Lara, A. C., Larson, A., & Losos, J. B. (2004). Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, *431*(7005), 177-181.
Li-Byarlay, H., Rittschof, C. C., Massey, J. H., Pittendrigh, B. R., & Robinson, G. E. (2014). Socially responsive effects of brain oxidative metabolism on aggression. *Proceedings of the National Academy of Sciences*, *111*(34), 12533-12537.
Linck, E., & Battey, C. (2019). Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Molecular Ecology Resources*, *19*(3), 639-647.
Locke, B. (2016). Natural Varroa mite-surviving Apis mellifera honeybee populations. *Apidologie*, *47*, 467-482.
Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in ecology & evolution*, *20*(5), 229-237.
Mattila, H. R., & Seeley, T. D. (2007). Genetic diversity in honey bee colonies enhances productivity and fitness. *Science*, *317*(5836), 362-364.
McNally, L., & Schneider, S. (1992). Seasonal cycles of growth, development and movement of the African honey bee, Apis mellifera scutettata, in Africa. *Insectes Sociaux*, *39*, 167-179.

McQuibban, G. A., Lee, J. R., Zheng, L., Juusola, M., & Freeman, M. (2006). Normal mitochondrial dynamics requires rhomboid-7 and affects Drosophila lifespan and neuronal function. *Current Biology*, *16*(10), 982-989.

Miguel, I., Baylac, M., Iriondo, M., Manzano, C., Garnery, L., & Estonba, A. (2011). Both geometric morphometric and microsatellite data consistently support the differentiation of the Apis mellifera M evolutionary branch. *Apidologie*, *42*, 150-161.

Moehring, A. J., Teeter, K. C., & Noor, M. A. (2007). Genome-wide patterns of expression in Drosophila pure species and hybrid males. II. Examination of multiple-species hybridizations, platforms, and life cycle stages. *Molecular biology and evolution*, *24*(1), 137-145.

Moorjani, P., & Hellenthal, G. (2023). Methods for Assessing Population Relationships and History Using Genomic Data. *Annual review of genomics and human genetics*, *24*(1), 305-332.

Moorjani, P., Patterson, N., Hirschhorn, J. N., Keinan, A., Hao, L., Atzmon, G., Burns, E., Ostrer, H., Price, A. L., & Reich, D. (2011). The history of African gene flow into Southern Europeans, Levantines, and Jews. *PLOS Genetics*, *7*(4), e1001373.

Nelson, R. M., Wallberg, A., Simões, Z. L. P., Lawson, D. J., & Webster, M. T. (2017). Genomewide analysis of admixture and adaptation in the Africanized honeybee. *Molecular Ecology*, *26*(14), 3603-3617. https://doi.org/10.1111/mec.14122

Noël, A., Le Conte, Y., & Mondet, F. (2020). Varroa destructor: how does it harm Apis mellifera honey bees and what can be done about it? *Emerging Topics in Life Sciences*, *4*(1), 45-57.

Noor, M. A. (2005). Patterns of evolution of genes disrupted in expression in Drosophila species hybrids. *Genetics Research*, *85*(2), 119-125.

Noriega, F. G. (2014). Juvenile hormone biosynthesis in insects: what is new, what do we know, and what questions remain? *International scholarly research notices*, *2014*(1), 967361.

Picard Toolkit. In. (2019). [Picard2019toolkit]. Broad Institute.

Rand, E. E. d., Smit, S., Beukes, M., Apostolides, Z., Pirk, C. W., & Nicolson, S. W. (2015). Detoxification mechanisms of honey bees (Apis mellifera) resulting in tolerance of dietary nicotine. *Scientific Reports*, *5*(1), 11779.

Roffet-Salque, M., Regert, M., Evershed, R. P., Outram, A. K., Cramp, L. J., Decavallas, O., Dunne, J., Gerbault, P., Mileto, S., & Mirabaud, S. (2015). Widespread exploitation of the honeybee by early Neolithic farmers. *Nature*, *527*(7577), 226-230.

Ruttner, F. (1988). Morphometric Analysis and Classification. In *Biogeography and taxonomy of honeybees* (pp. 66-78). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-72649-1_6

Ruttner, F., & Ruttner, F. (1988). Morphometric analysis and classification. *Biogeography and taxonomy of honeybees*, 66-78.

Ruttner, F., Tassencourt, L., & Louveaux, J. (1978). Biometrical-statistical analysis of the geographic variability of Apis mellifera LI Material and methods. *Apidologie*, *9*(4), 363-381.

Saravanan, K., Panigrahi, M., Kumar, H., Bhushan, B., Dutt, T., & Mishra, B. (2020). Selection signatures in livestock genome: A review of concepts, approaches and applications. *Livestock Science*, *241*, 104257.

Schmidt, T. L., Jasper, M. E., Weeks, A. R., & Hoffmann, A. A. (2021). Unbiased population heterozygosity estimates from genome-wide sequence data. *Methods in Ecology and Evolution*, *12*(10), 1888-1898.

Scott Schneider, S., DeGrandi-Hoffman, G., & Smith, D. R. (2004). The African honey bee: factors contributing to a successful biological invasion. *Annual Reviews in Entomology*, *49*(1), 351-376.

Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic acids research*, *50*(W1), W216-W221.

Spötter, A., Gupta, P., Nürnberg, G., Reinsch, N., & Bienefeld, K. (2012). Development of a 44K SNP assay focussing on the analysis of a varroa-specific defence behaviour in honey bees (Apis mellifera carnica). *Molecular Ecology Resources*, *12*(2), 323-332.

Tarpy, D. R. (2003). Genetic diversity within honeybee colonies prevents severe infections and promotes colony growth. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1510), 99-103.

Tonks, N. K. (2006). Protein tyrosine phosphatases: from genes, to function, to disease. *Nature reviews Molecular cell biology*, *7*(11), 833-846.

Toth, A. L., & Zayed, A. (2021). The honey bee genome-- what has it been good for? *Apidologie*, *52*(1), 45-62. https://doi.org/10.1007/s13592-020-00829-3

True, J. R. (2003). Insect melanism: the molecules matter. *Trends in ecology & evolution*, *18*(12), 640-647.

Tubaro, P. L., & Lijtmaer, D. A. (2002). Hybridization patterns and the evolution of reproductive isolation in ducks. *Biological Journal of the Linnean Society*, *77*(2), 193-200.

Tzec-Interián, J. A., González-Padilla, D., & Góngora-Castillo, E. B. (2025). Bioinformatics perspectives on transcriptomics: A comprehensive review of bulk and single-cell RNA sequencing analyses. *Quantitative Biology*, *13*(2), e78.

Wallberg, A., Bunikis, I., Pettersson, O. V., Mosbech, M.-B., Childers, A. K., Evans, J. D., Mikheyev, A. S., Robertson, H. M., Robinson, G. E., & Webster, M. T. (2019). A hybrid de novo genome assembly of the honeybee, Apis mellifera, with chromosome-length scaffolds. *BMC Genomics*, *20*(1), 275. https://doi.org/10.1186/s12864-019-5642-0

Wallberg, A., Glémin, S., & Webster, M. T. (2015). Extreme recombination frequencies shape genome variation and evolution in the honeybee, Apis mellifera. *PLOS Genetics*, *11*(4), e1005189.

Wallberg, A., Han, F., Wellhagen, G., Dahle, B., Kawata, M., Haddad, N., Simões, Z. L. P., Allsopp, M. H., Kandemir, I., & De la Rúa, P. (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.

Walsh, A. T., Triant, D. A., Le Tourneau, J. J., Shamimuzzaman, M., & Elsik, C. G. (2022). Hymenoptera Genome Database: new genomes and annotation datasets for improved go enrichment and orthologue analyses. *Nucleic acids research*, *50*(D1), D1032-D1039.

Wang, K., Liu, Z.-g., Lin, Z.-g., Yin, L., Gao, F.-c., Chen, G.-h., & Ji, T. (2020). Epigenetic modifications may regulate the activation of the hypopharyngeal gland of honeybees (Apis mellifera) during winter. *Frontiers in Genetics*, *11*, 46.

Whitfield, C. W., Behura, S. K., Berlocher, S. H., Clark, A. G., Johnston, J. S., Sheppard, W. S., Smith, D. R., Suarez, A. V., Weaver, D., & Tsutsui, N. D. (2006). Thrice Out of Africa: Ancient and Recent Expansions of the Honey Bee, Apis mellifera. *Science*, *314*(5799), 642-645. https://doi.org/doi:10.1126/science.1132772

Winston, M. L. (1992). *Killer bees: the Africanized honey bee in the Americas*. Harvard University Press.

Wittkopp, P. J., & Beldade, P. (2009). Development and evolution of insect pigmentation: genetic mechanisms and the potential consequences of pleiotropy. Seminars in cell & developmental biology,

Zanni, V., Galbraith, D. A., Annoscia, D., Grozinger, C. M., & Nazzi, F. (2017). Transcriptional signatures of parasitization and markers of colony decline in Varroa-infested honey bees (Apis mellifera). *Insect biochemistry and molecular biology*, *87*, 1-13.

Zheng, X., Gogarten, S. M., Lawrence, M., Stilp, A., Conomos, M. P., Weir, B. S., Laurie, C., & Levine, D. (2017). SeqArray—a storage-efficient high-performance data format for WGS variant calls. *Bioinformatics*, *33*(15), 2251-2257.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, *28*(24), 3326-3328.