Hindawi
Psyche: A Journal of Entomology
Volume 2024, Article ID 6627723, 7 pages
https://doi.org/10.1155/2024/6627723

Research Article

Genome Resequencing of the Honeybee *Apis mellifera jemenetica* (Hymenoptera: Apidae): A Key Tool towards Characterization, Conservation, and Genomic Selection

Ahmad Alghamdi and Yehya Alattal

Department of Plant Protection, Chair of Engineer Abdullah Ahmad Bagshan for Bee Research, College of Food and Agriculture Sciences, King Saud University, Riyadh 11451, Saudi Arabia

Correspondence should be addressed to Yehya Alattal; yalattal@ksu.edu.sa

Received 16 June 2023; Revised 18 August 2023; Accepted 28 December 2023; Published 16 January 2024

Academic Editor: Cleber Galvão

We report the whole-genome sequence of the Arabian honeybee (*Apis mellifera jemenetica*). Seven *A. m. jemenetica* samples were sequenced representing three distinct subpopulations. Generated sequence reads were mapped to the reference honeybee *Apis mellifera* genome (Amel_HAv3.1). Data revealed genome-wide patterns of genetic variation which can be useful in the characterization and assessment of positive selection of the Arabian honeybee using different genetic markers. In total, 75.16 Gb of clean bases were generated, and the GC content of samples ranged between 31.9 and 35.3%. The effective reference genome size is 223,937,270 bp. The mapping rate of samples varied from 88.97% to 96.19%, and the effective mapping depth was between 41.80 and 48.84X. Single-nucleotide polymorphisms (SNPs) among sequenced individuals ranged between 2379499 and 2396116 with respect to the reference *A. mellifera* genome (Amel_HAv3.1), and 2% of the SNPs were nonsynonymous. Genome Analysis Toolkit (GATK) detected 1097962–1109829 InDels and 10090–11962 structural variations (SV) from which 22.1 to 33.8% were in the form of deletions. Copy number variation (CNV) ranged between 550 and 2824, and 45–91% of them were downregulated. These variations among interbreeding individuals or groups of the same species may reflect an adaptive environmental response and fitness among different subpopulations and can be very useful for subspecies characterization, conservation, and selection of the Arabian honeybee.

1. Introduction

The honeybee *Apis mellifera* (Linnaeus, 1758) (Hymenoptera: Apidae) comprises 31 geographical subspecies distributed in Asia, Africa, and Europe [1]. Based on molecular and/or morphometric analyses, these subspecies have been classified into separate evolutionary lineages and sublineages: the O and Y lineages occurring in Western Asia, the M lineage in Eurasia, the C lineage in Europe, and the A lineage and Z sublineage in Africa [1, 2]. The Arabian honeybee *A. m. jemenetica* is the native honeybee subspecies of the Arabian Peninsula and tropical Africa. It is the only *A. mellifera* subspecies that is naturally occurring in Africa and Asia, exhibiting distinctive traits associated with different geographical ranges [2]. It is assumed that beekeeping in Saudi Arabia started since for more than 4000 years ago [3]. *A. m. jemenetica* has been kept in diverse climatic zones of the country and acquired unique morphological and behavioral aspects among other *A. mellifera* subspecies. For example, it is the smallest and the most adapted to extreme temperatures compared to other European or African *A. mellifera* subspecies [2, 4, 5]. Each honeybee lineage or sublineage can include several *A. mellifera* subspecies. Based on morphometric characterization and mitochondrial DNA analysis, *A. m. jemenetica* of Saudi Arabia belongs to the Y lineage (previously called Z sublineage or O lineage) and is separated into 3 distinctive subpopulations, with significant morphological/molecular variations [6–8]. Recently, a comprehensive phylogenetic analysis of the Arabian honeybee, *A. m. jemenetica*, based on complete

mitochondrial DNA sequences ($n = 14$), revealed three distinct subclusters within Saudi Arabia [8], with very close evolutionary relationship with the Syrian honeybee, *A. m. syriaca*, the Egyptian honeybee *A. m. lamarckii*, and east African subspecies such as *A. m. capensis* and *A. m. scutellata* [8]. These subpopulations may exhibit different adaptive traits [9]. Nevertheless, huge importations of exotic honeybee subspecies may describe Saudi Arabia as the foremost package bee importer worldwide with about 1.3 million packages being imported in 2021 [10], which may impact the presence and characteristics of *A. m. jemenetica* and entails urgent conservation strategies of this honeybee subspecies.

The completion of the honeybee genome (*A. m. ligustica*) [11] and subsequent releases and updates opens the door for genome-wise comparisons with other *A. mellifera* subspecies and ecotypes. From the 31 reported *A. mellifera* subspecies, only a few *A. mellifera* subspecies genomes were published (*A. m. ligustica*, *A. m. carnica*, *A. m. mellifera*, *A. m. caucasica*, and *A. m. intermissa*). Recently, by analyzing 251 genomes of 18 native honeybee subspecies, scientists suggested a West Asian origin of *A. mellifera* with at least three adaptive radiations given to African and European lineages; consequently, 145 genes with unique signatures of selection across all bee lineages were documented [12]. These findings increased the focus on the characterization and conservation of the West Asian honeybee subspecies such as *A. m. jemenetica* and *A. m. syriaca*. In this study, the whole-genome sequences of seven honeybee samples representing three distinct subpopulations of the Arabian honeybee *A. m. jemenetica* have been reported, which will enrich our understanding of the evolutionary relationship among *A. mellifera* subspecies and provide a wealth of knowledge on molecular characterization, conservation, and selection of *A. m. jemenetica* [13].

2. Materials and Methods

2.1. Sampling and DNA Extraction. Seven native nonmigratory honeybee apiaries (location (sample code): Gis): Al-Madina (MD101.1): 25.411561, 37.520384; Najran (N21): 17.625917, 43.754611; Asir (A43) 18.256722, 42.229028; Albaha (B31): 19.852472, 41.585611; Makkah (MK163): 21.911416, 39.753706; Jazan (J61): 17.517472, 43.075000; and Tabuk (Tu21): 28.387028, 36.855194 were selected from purebred beekeeping areas within Saudi Arabia. From each apiary, one sample consisting of 15 bees was collected, 10 bees were characterized morphometrically using body dimensions (proboscis, femur, tibia, metatarsus, hind leg, tergite-3, and tergite-4 cubital index 1 and 2) and wing characters (wing angles: a4, b4, d7, c9, g18, j10, j16, k19, l13, n23, and o26 and 20 wing landmarks) [2, 14] to ensure accurate subspecies affiliation. After that, three adult workers were taken from each sample. Bee samples were preserved in ethanol (96%) and stored in the freezer. Genomic DNA was extracted from one honeybee and purified using Qiagen extraction and purification kits (CAT: 96504; 28106, respectively: Qiagen, Hilden, Germany). The genomic DNA was then sent to BGI (https://www.bgi.com) for sequencing. The genomic DNA was first subjected to ultrasound shearing, end repairing, and end adaptor ligation and fragment selection for amplification and library construction.

2.2. Sequencing Assessment and Raw Data Processing. Figure 1 charts the steps of sequencer raw data processing, sequence alignment, and sequence variation assessments. Raw sequencing data produced from the DNBseq-G400 pine line were subjected to three-step filtration using SOAPnuke v1.5.6 software [15]. Data processing started with adaptor trimming, and any reads with an adaptor mapping rate higher than 50% were removed. Then, low-quality reads with more than 50% of low-quality bases ($Q20 < 50\%$) were removed. Finally, contiguous reads with more than 2% N bases were removed.

2.3. Alignment. Sequence reads were aligned to the reference genome using the BWA analysis tool (https://bio-bwa.sourceforge.net/bwa.shtml) [11]. The honeybee Amel_HAv3.1 (https://www.ncbi.nlm.nih.gov/assembly/gcf_003254395.5) reference genome was used for alignment of sample sequence reads. The reference genome size is 225,250,884 bp while the effective size is 223,937,270 bp (N base excluded), with GC content at 32.34%. The mapping rate of samples varied from 32.96% to 99.47%, and the effective mapping depth ranged between 26.51X and 59.33X (according to the sample selected). The short-sequence reads aligned against the long reference genome were created in SAM (Sequence Alignment/Map) format. Picard tools (v1.118) were used to sort the SAM files by coordinate and convert them to BAM files. Picard tools software (v1.118) (https://broadinstitute.github.io/picard/) was also used to mark duplicate reads.

2.4. Sequence Variation Assessment. GATK (version 3.7-0-gcfedb67: https://www.broadinstitute.org/gatk/, RealnTimes = 1) was used in single-nucleotide polymorphism (SNPs) and short insertions and deletions (InDels) calling. SNPs annotation started with a consensus sequence which consisted of all bases of the target genome. SNPs that existed in both genotypes (our genotype and the reference genomes) were screened, and after that, a highly reliable SNP dataset was generated. Nonsynonymous mutation and large-effect SNPs were also recorded. Annotation and statistics of selected SNPs and InDels BGI were performed using an in-house Perl script and ReSeqTools (https://github.com/BGI-shenzhen/Reseqtools). The BreakDancer/CREST method (https://breakdancer.sourceforge.net [11]). Based on reference genome, the BreakDncer/CREST method (https//breakdancer.sourceforge.net (chen et al., 2009)) enabled us to produce a list of structural variations (SV) that were detected across the whole genome. Copy number variations (CNVs) were detected according to Zheng et al. [16] method. Using SOAP alignment results, the depth of each base was calculated and standardized by the mean depth of its chromosome to calculate the copy number variation according to the standard method [16].

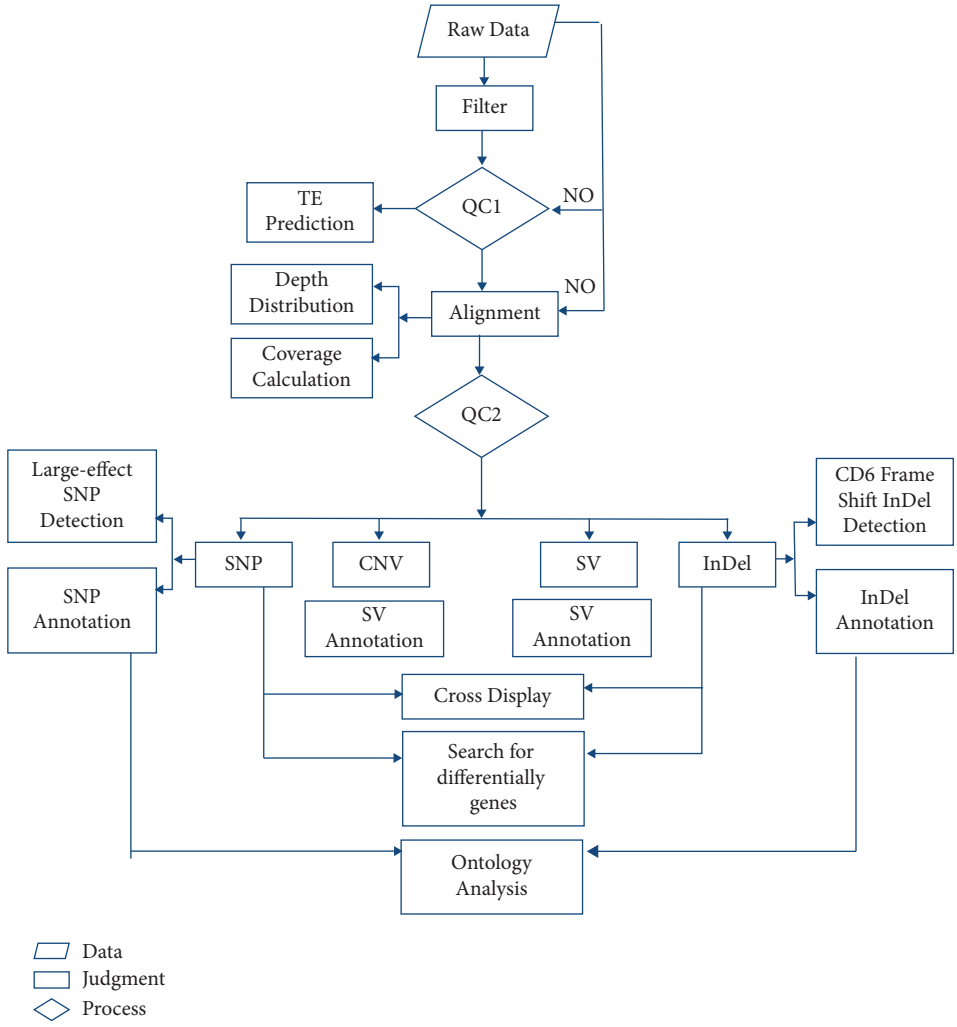

Figure 1: Graph of standard bioinformatic raw data cleaning and analysis based on the family SOAP software (TE: transposable elements; QC: quality control; CNV: copy number variation; SV: structural variation).

3. Results

3.1. Sequences and Alignment Assessment. In total, 75.16 G of clean bases were generated. About 96% of the clean data have quality values higher than 20 ($q20 \geq 96\%$) and from 85.2 to 90.4% of the data have quality values higher than 30 ($q30 \geq 96\%$) (Table 1). For alignment, the *Apis mellifera ligustica* reference genome (assemblyAmel_HAv3.1: https://www.ncbi.nlm.nih.gov/genome/48?genome_assembly_id=403979) was used. The reference genome size is 225,250,884 bp, while the effective size is 223,937,270 bp (N base excluded), with GC content at 32.34%. The mapping rate of samples varies from 88.97% to 96.19%, and the effective mapping depth is between 41.80 and 48.84X (Table 2). GC content of samples ranged between 31.9 and 35.3 (Table 1).

3.2. SNPs Detection and Annotation. All seven samples represented morphologically one honeybee subspecies *A. m. jemenetica* and three distinct subpopulations. The polymorphic loci in *A. m. jemenetica* samples revealed a high total number of SNPs ranging between 2379499 and 2396116 in relation to the reference *A. mellifera* genome (Amel_HAv3.1) and 35.1–36.3% of which were heterozygous. Nonsynonymous SNPs counted about 2% of the total number of SPNs and ranged between 48892 and 45559. Polymorphic sites of exons ranged between 529533 and 540719. A total number of 2041833–2050427 genes had at least one polymorphic site among different honeybee samples. Sample MD101 showed the highest number of nonsynonymous SNPs and SNPs on exons. These variations may have a great impact on the biological or behavioral traits of the Arabian honeybee *A. m. jemenetica*. The total number and distribution of SNPs are listed in Table 3.

3.3. Insertions and Deletions (InDels). Genome Analysis Toolkit (GATK) detected 1097962–1109829 insertions and deletions (InDels) among different *A. m. jemenetica* samples, which were distributed almost equally between insertions and deletions. Table 4 shows the distribution and annotation of InDels. Insertion and deletions at coding regions (CDS-InDels) were 0.94–0.99 of the total InDels in the genomes of different samples. About 0.86% of the InDels occurred within genes. InDel annotation can reveal more details about the evolutionary history among *A. mellifera* subspecies.

Table 1: Quality evaluation of clean reads data of different Arabian honeybee *A. m. jemenetica* sequences representing three subpopulations.

Sample	GC rate (%)	Q20_rate (%)	Q30_rate (%)	Bases (G)	CDR
J61 (SRR12778250*)	32.38	96.85	88.72	09.42	99.47
B31 (SRR12778251)	34.73	96.13	86.72	10.72	99.63
A43 (SRR12778252)	32.19	96.24	86.8	11.01	99.55
MK163 (SRR12778253)	33.32	96.32	87.21	10.93	98.82
N21 (SRR12778254)	33.35	96.89	88.94	11.15	98.67
TU21 (SRR12778255)	31.91	95.86	85.2	11.06	99.46
MD101 (SRR12778256)	35.33	97.49	90.4	10.94	98.92

*Accession number.

Table 2: Alignment analysis of the clean read sequences of the Arabian honeybee *A. m. jemenetica* mapped to the reference *A. m. jemenetica* genome (Amel_HAv3.1).

Sample	Coverage rate (%)	Map_reads_rate (%)	Map_bases_rate (%)	Uni-hit-reads rate (%)	Uni-hit-bases rate (%)	S depth
J61	98.43	94.00	94.00	93.12	87.53	41.80
B31	98.27	92.70	92.70	92.90	86.12	47.58
A43	98.47	94.63	94.63	92.81	87.83	48.86
MK163	98.45	94.36	94.36	92.19	86.99	48.52
N21	98.44	88.97	88.97	92.45	82.25	49.48
TU21	98.49	96.19	96.19	92.23	88.72	48.84
MD101	97.23	97.42	97.42	92.82	90.42	48.56

Table 3: Genome-wide distribution and annotation of single-nucleotide polymorphisms (SNPs) in the Arabian honeybee samples.

Annotation	Sample						
	J61	B31	A43	MK163	N21	TU21	MD101
SNPs	2393148	2382428	2396116	2393932	2390554	2394339	2379499
Homozygous	1548264	1547992	1555807	1549604	1551806	1555838	1514837
Heterozygous	844884	834436	840309	844328	838748	838501	864662
Synonymous	177421	178610	181644	178532	185311	182409	186954
Nonsynonymous	45559	47037	45949	47754	47178	46412	48892
Exons	537255	533837	535517	535890	535990	529533	540716
Genes	2048575	2041833	2050427	2048995	2045289	2045994	2044125
mRNA	30	29	25	34	29	27	33
Pseudogene	8041235	8003125	8047514	8048893	8030989	8010708	8031140
rRNA	716	688	739	759	748	777	718
snoRNA	25	29	24	27	25	23	31
siRNA	8	9	7	9	9	9	11
snRNA	13	12	14	14	10	13	11
Transcript	311994	311501	312475	310577	311631	307971	310887
tRNA	19	14	13	16	16	13	16

3.4. SV and CNV. Structural variation is a large sequence variation (1 kb or even larger) that can include sequence duplication, inversion, and transposition or insertions and deletions, commonly referred to as copy number variants (CNVs). These CNVs often overlap with segmental duplications, and regions of DNA>1 kb exist more than once in the genome, copies of which are >90% identical [17]. If present at >1% in a population, a CNV may be referred to as copy number polymorphism (CNP). The highest number of structural variations (SVs) and copy number variations occurred in the samples MD101 and B31. The copy number variation in the MD101 genome was about 4 times more compared with other *A. m. jemenetica* genomes. About 22.1 to 33.8 of the structural variation in the samples' sequences were in the form of deletion, and Table 5 lists the detailed distribution and annotation of the structural variations in all *A. m. jemenetica* samples. Copy number variations among individuals of the same species may reflect environmental response, for example, sensory perception and immunity [18, 19], and is considered an impotent genetic variation among populations. About 45–91% of copy number variations were downregulated. Table 6 lists the detailed distribution and annotation of the copy number variation in the honeybee *A. mellifera*.

Table 4: Genome-wide distribution and annotation of insertions and deletions (InDel) in the Arabian honeybee samples.

Annotation	Sample						
	J61	B31	A43	MK163	N21	TU21	MD101
InDels	1107472	1099320	1109829	1109000	1105877	1108522	1097962
Insertion	547740	544015	549648	548898	547293	549025	543466
Deletion	559732	555305	560181	560102	558584	559497	554496
CDS	10783	10363	10836	10735	10712	10690	10880
cDNA_match	1168	1132	1159	1150	1182	1141	1151
Exon	170666	169932	172038	171654	170517	171749	169178
Gene	952464	947684	954046	953032	950553	951769	943518
Inc-RNA	167712	167121	168183	165626	167301	167950	165486
mRNA	3720850	3699218	3724766	3723011	3712757	3709350	3685324
miRNA	14	11	11	12	16	10	11
Primary_transcript	45	43	37	42	48	37	39
Pseudogene	254	231	257	248	257	272	243
rRNA	6	8	9	11	12	11	10
Region	1107472	1099320	1109829	1109000	1105877	1108522	1096162
Sequence feature	2	1	1	2	7	5	1
SnoRNA	7	11	11	11	12	14	10
tRNA	118826	118706	118229	117953	118539	116995	117069
Transcript	0	**0**	**0**	0	**0**	0	0

Table 5: Genome-wide distribution and annotation of structural variation (SV) in the Arabian honeybee samples.

Annotation	Sample						
	J61	B31	A43	MK163	N21	TU21	MD101
SV	11088	10090	10635	11111	10854	11178	11962
Deletion	2718	3389	3596	2624	2592	3413	2648
CDS	37155	13808	40932	11733	29185	28568	17763
cDNA_match	183	79	278	97	214	142	109
Exon	44518	17119	49293	14317	34910	34941	21972
Gene	11349	9391	11335	9967	10787	11114	11196
Inc-RNA	2041	1725	2043	1607	1870	2000	1877
mRNA	37381	33522	34616	36285	35047	36201	40919
miRNA	55	27	42	18	40	45	31
Primary_transcript	54	27	42	18	40	45	31
Pseudogene	12	12	15	9	12	12	10
rRNA	5	20	10	11	14	13	10
Region	11087	10090	10634	11111	10853	11178	11962
Sequence feature	2	2	3	11	2	2	2
snRNA	3	1	2		1	1	2
SnoRNA	2	9	13	0	29	18	12
tRNA	23	1021	1180	1185	1233	1310	1367
Transcript	1269	0	0	**0**	0	0	0

4. Discussion

One of the striking features of *A. mellifera* is the high recombination between chromosomes, which favors effective natural selection. Data presented here are the whole-genome resequencing of the native honeybee subspecies of Saudi Arabia *A. m. jemenetica*. Results demonstrated high genetic variation in SNPs, SV, InDels, and CVN among *A. m. jemenetica* samples and in comparison to the reference *A. mellifera* genome sequence. These genetic variations can be used in the molecular characterization of *A. m. jemenetica* and related subpopulations [20]. Furthermore, adaptive structural and behavioral traits associated with the natural occurrence of *A. m. jemenetica* can be explored when data are subjected to further analysis [11].

The very high CNV in two samples (MD101 and B31) can be related to nonallelic homologous recombination during meiotic recombination and may demonstrate direct interaction with different environments [19]. Furthermore, high GC content in some genomes of the native honeybee of Saudi Arabia may indicate higher recombination rates [11] or higher genetic diversity [11]. Single-nucleotide polymorphisms (SNPs) are variations caused by the changing of a single nucleotide (A, T, C, or G) in the genome. The SNPs, including switch and reverse of single-nucleotide bases, are responsible for genome diversity between species and among individuals of the sample species. Analysis of SNPs among honeybee subspecies across genomes has been used in testing hypotheses concerning the ancestral origin of the western honeybee

Table 6: Genome-wide distribution and annotation of copy number variation (CNV) in the Arabian honeybee samples.

Annotation	Sample						
	J61	B31	A43	MK163	N21	TU21	MD101
CNV	552	2113	552	550	553	524	2824
Upregulated	230	225	207	201	208	212	250
Downregulated	322	1888	345	249	345	312	2574
CDS	473	6893	310	574	449	601	12694
cDNA_match	14	56	12	11	9	10	115
Exon	1986	9921	1808	2057	1914	1856	16588
Gene	268	1968	274	281	285	252	3082
Inc-RNA	335	435	329	331	275	297	443
mRNA	969	5288	921	1002	971	800	7673
miRNA	2	5	0	1	2	0	2
Primary_transcript	2	5	0	1	2	0	4
Pseudogene	12	14	11	11	11	8	11
rRNA	5	10	9	7	21	10	6
Region	552	2113	552	550	601	524	2824
Sequence feature	1	0	0	0	0	0	0
snRNA	0	0	0	0	0	0	1
SnoRNA	0	1	0	0	0	0	1
tRNA	1	3	1	0	1	0	28
Transcript	67	230	60	73	100	44	327

A. mellifera and the relationship among a large number of its subspecies [2]. SNP data created from this publication may help in further understanding of *A. m. jemenetica* evolution and natural selection. Additionally, a positive correlation between SNP number and GC content can indicate many remarkable evolutionary features [11] and molecular adaptive characteristics [11]. Previous studies indicated that the GC-rich regions have more mutations [11], and this may be the case for the native honeybee populations in some regions of Saudi Arabia. InDels refers to insertion mutation, deletion mutation, or both, including events in the early stage of evolution, and might help in explaining levels of gene expression that disrupt the production of essential proteins [21]. Copy number variation (CNV) is an important form of structural variation among individuals of the same species which may indicate long-term adaption to the unique environment of the Arabian Peninsula; CNVs of genes coding to vitellogenin or heat shock proteins or drought tolerance may unveil molecular aspects of adaptation of this honeybee subspecies compared to others.

5. Conclusion

This is the first genome resequencing of the Arabian honeybee *A. m. jemenetica*. A honeybee that is characterized as unique in many aspects (size, heat tolerance, and natural distribution). Whereas the complete mitochondrial analysis of the Arabian honeybee from Saudi Arabia focuses on subspecies affiliation [8], the whole-genome sequences focus on in-depth genetic variation among the Arabian honeybee populations within Saudi Arabia. The outcome of this study can be used in the characterization, selection, and breeding of *A. m. jemenetica*.

Data Availability

The genomes have been uploaded to the National Center for Biotechnology Information (NCBI: Sequence Reads Archives) and were assigned the following accession numbers: https://www.ncbi.nlm.nih.gov/search/all/?term=SAMN16378534 to SAMN16378541.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was funded by the National Plan for Science, Technology, and Innovation (MAARIFAH), King Abdul-Aziz City for Science and Technology, Kingdom of Saudi Arabia, grant no. 13-AGR1443-02. We are grateful to local beekeepers who allowed us to sample their apiaries and to Iftikhar Rasool and Arif Single in morphological slide preparation.

References

[1] R. A. Ilyasov, M. Lee, J. Takahashi, H. W. Kwon, and A. G. Nikolenko, "A revision of subspecies structure of western honeybee *Apis mellifera*," *Saudi Journal of Biological Sciences*, vol. 27, pp. 3615–3621, 2020.

[2] F. Ruttner, *Biogeography and Taxonomy of Honeybees*, Springer, Berlin, Germany, 1st edition, 1988.

[3] M. Engel, A. Alqarni, M. Hannan, and A. Owayss, "The indigenous honey bees of Saudi Arabia (Hymenoptera, Apidae, *Apis mellifera jemenitica* Ruttner): their natural history and role in beekeeping," *ZooKeys*, vol. 134, pp. 83–98, 2011.

[4] Y. Alattal and A. Alghamdi, "Impact of temperature extremes on survival of indigenous and exotic honey bee subspecies, *Apis mellifera*, under desert and semiarid climates," *Bulletin of Insectology*, vol. 68, pp. 219–222, 2015.

[5] Y. Z. Alattal and A. A. Alghamdi, "Linking histone methylation states and hsp transcriptional regulation in thermotolerant and thermo-susceptible *A. mellifera* L. Subspecies in response to heat stress," *Insects*, vol. 14, no. 3, p. 225, 2023.

[6] A. Alghamdi, M. Alsharhi, Y. Alattal, and N. Adgaba, "Morphometric diversity of indigenous honeybees, *Apis mellifera* (Linnaeus, 1758), in Saudi Arabia," *Zoology in the Middle East*, vol. 58, pp. 97–103, 2012.

[7] Y. Alattal, A. Alghamdi, M. Alsharhi, and S. Fuchs, "Morphometric characterization of the native honeybee, *Apis mellifera* Linnaeus, 1758, of Saudi Arabia," *Zoology in the Middle East*, vol. 60, no. 3, pp. 226–235, 2014.

[8] A. Alghamdi and Y. Alattal, "The complete mitochondrial genome and phylogenetic analysis of the Arabian Honeybee *Apis mellifera jemenitica* (Insecta: hymenoptera: Apidae) from Saudi Arabia," *Mitochondrial Deoxyribonucleic acid Part B*, vol. 5, no. 3, pp. 3655–3657, 2020.

[9] Y. Alattal and A. Alghamdi, "Evidence for sub-populations of *Apis mellifera jemenitica* colonies along the Red Sea coast of Saudi Arabia," *Bulletin of Insectology*, vol. 71, pp. 7–14, 2022.

[10] MoEP Import, "Data of honeybees from different sources," in *Central Department of Statistics and information*, Ministry of Economy and Planning, Kingdom of Saudi Arabia, 2022.

[11] The Honeybee Genome Sequencing Consortium, "Insights into social insects from the genome of the honeybee *Apis mellifera*," *Nature*, vol. 443, no. 7114, pp. 931–949, 2006.

[12] K. A. Dogantzis, T. Tiwari, I. M. Conflitti et al., "Thrice out of Asia and the adaptive radiation of the western honey bee," *Science Advances*, vol. 7, no. 49, Article ID eabj2151, 2021.

[13] B. A. Harpur, C. F. Kent, D. Molodtsova et al., "Population genomics of the honey bee reveals strong signatures of positive selection on worker traits," *Proceedings of the National Academy of Sciences*, vol. 111, no. 7, pp. 2614–2619, 2014.

[14] I. Kandemir, M. G. Moradi, B. Özden, and A. Özkan, "Wing geometry as a tool for studying the population structure of dwarf honey bees (Apis florea Fabricius 1876) in Iran," *Journal of Apicultural Research*, vol. 48, no. 4, pp. 238–246, 2009.

[15] Y. Chen, Y. Chen, C. Shi et al., "SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data," *GigaScience*, vol. 7, no. 1, pp. 1–6, 2018.

[16] L. Y. Zheng, X. S. Guo, B. He et al., "Genome-wide patterns of genetic variation in sweet and grain sorghum (Sorghum bicolor)," *Genome Biology*, vol. 12, no. 11, p. R114, 2011.

[17] A. J. Sharp, D. P. Locke, S. D. McGrath et al., "Segmental duplications and copy-number variation in the human genome," *The American Journal of Human Genetics*, vol. 77, no. 1, pp. 78–88, 2005.

[18] D. F. Conrad, T. D. Andrews, N. P. Carter, M. E. Hurles, and J. K. Pritchard, "A high-resolution survey of deletion polymorphism in the human genome," *Nature Genetics*, vol. 38, no. 1, pp. 75–81, 2006.

[19] R. Redon, S. Ishikawa, K. Fitch et al., "Global variation in copy number in the human genome," *Nature*, vol. 444, no. 7118, pp. 444–454, 2006.

[20] J. Momeni, M. Parejo, R. O. Nielsen et al., "Authoritative subspecies diagnosis tool for European honey bees based on ancestry informative SNPs," *Bone Marrow Concentrate Genomics*, vol. 22, no. 1, p. 101, 2021.

[21] W. Sung, M. S. Ackerman, M. M. Dillon et al., "Evolution of the insertion-deletion mutation rate across the tree of life," *G3 Genes Genomes GeneticsG3 Genes Genomes Genetics*, vol. 6, no. 8, pp. 2583–2591, 2016.