First record of the complete mitochondrial genome of *Culex pipiens* (Diptera: Culicidae) in Iraq and its phylogenetic analysis

Zhuang-Xin Ye[1], Adnan A. Lahuf [2], Mohammed D. Salman[2], Jun-Min Li[1]

[1] State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-products, Key Laboratory of Biotechnology in Plant Protection of Ministry of Agriculture and Zhejiang Province, Institute of Plant Virology, Ningbo University, Ningbo, China
[2] Plant Protection Department, Agriculture College, University of Kerbala, Karbala, Iraq
***Corresponding author e-mail: adnan.lahuf@uokerbala.edu.iq**

Received:
Jan. 27, 2023

Accepted:
Mar. 01, 2023

Published:
Mar. 23, 2023

Abstract

The common house mosquito, Culex pipiens, is a worldwide insect that is known as a vector of several human and animals' pathogens that cause critical diseases, such as malaria, yellow fever and encephalitis. The mitochondrial genome knowledge can supply a foresight for evolutionary biology and phylogenetic analysis. Thus, in the current study, the complete mitochondrial genomes of the Iraqi C. pipiens were first sequenced, assembled, annotated, and analyzed using high throughput sequencing techniques. The results revealed that the length of the complete mitochondrial genome is 14,856 bp, comprising 13 protein-coding genes (PCGs), 22 transfer RNA genes, and two ribosomal RNA genes. The A + T content was 77.09%, while the G+C content was only 22.1%. All PCGs open with the start codon ATN, except the cox1 gene and end with the stop codon TAA. A phylogenetic relatedness with other different species of Culex spp. was achieved by operating the molecular evolutionary genetic analysis based on the whole mitochondrial genome sequences. The results showed that *C. pipiens* from Iraq shares a close ancestry with *C. pipiens* from Tunisia and France. To our knowledge, this is the first report of the complete mitochondrial genome sequence of the Iraqi C. pipiens. As a consequence, this conclusion may provide a new profound insight into the evolution of *C. pipiens* in Iraq.

Keywords: mitochondrial genome; *Culex pipiens*; phylogenetic analysis

Introduction

The mosquito, *Culex pipiens* (Culicidae: Diptera) is widely distributed [1]. Its global distribution in all tropical and temperate regions, reaches even the Arctic region. As well as, it was reported at an altitude of 5500 meters, and a depth of 1250 meters below sea level [2]. *C. pipiens* is one of the blood-sucking arthropods. It is also one of the most common types of mosquitoes known for their affinity for humans and animals, as they prefer environments containing high concentrations of organic matter [3]. Furthermore, it is a vector of many pathogens responsible for causing many serious diseases in humans and animals, such as human filarial nematodes, protozoa that causes

avian malaria, and several viruses causing yellow fever, rift valley fever and encephalitis [4; 5; 6].

Morphologically, eggs of mosquito demand water to hatch together (typically 150-300 eggs) as an ovoid bundle on the surface of still water. They remain floating until larval emergence. The larval stage is light brown in colour and small in size. Their head is wider than the rest of the body. They develop through four instars before transforming into active, non-feeding pupae. The adult's wings, sucking mouthparts, and legs can be seen through the transparent pupal skin; the adult emerges from the pupal stage to the water's surface [7]. The adult body is small to medium in size and light brown with scattered patches of pale scales. It comprises the head (eyes, antennae, proboscis, and palps), thorax (legs and wings), and abdomen (ten segments and genitalia) [8]. Generally, appropriate identification is indispensable for best control procedures [9]. Insects, including mosquitoes can be classified using their morphological characteristics, which profoundly depend on numerous exterior traits [10]. However, studies into *C. pipiens* have been very limited in their molecular description.

The mitochondrial (mt) genome codes conservation, maternal inheritance, quick evolution, and almost nonexistent intermolecular genetic recombination. Thus, it has become applicable and standard markers for molecular evolution and phylogenetic studies, such as phylogenetic molecular evolution, population genetics, and comparative evolutionary genomics [11; 12]. Usually, the mt genome is a closed circular arrangement with a length extended from nearly 14,000 to 18,000 bp. The outer circle is a heavy chain, while the inner circle is a light chain. It is a semi-autonomous organelle that can express its assured proteins separately in the cells. Due to the mtDNA being expressed as matrilineal inheritance, its structure and proportion can better reveal the genetic characteristics of populations, particularly for insects [13].

Hence, the object of the current study was to identify and classify the Iraqi mosquito, *Culex* sp. at the level of molecular biological evolution by sequencing the complete length of its mitochondrial genome.

Material and Methods

Sample Collection, nucleic acid extraction and next-generation sequencing

The common house mosquito *Culex* sp. samples were collected using aspirator traps from different locations of Karbala Province, Iraq, in 2020. The samples were placed in DNA/RNA Shield solution (ZYMO RESEARCH, Irvine, CA 92614, USA). The collected samples (5 adult mosquitoes per clean centrifugal tube) were then sent to Macrogen company (Seoul, South Korea). The nucleic acid extraction, the sequencing library construction and next-generation sequencing using Illumina NextSeq 1000 & 2000 platform were conducted according to the standard instructions of the company.

Assembly and Gene Annotation

In this study, the next-generation sequencing (NGS) raw reads data received were checked using FastQC software [14; 15], and De Novo assembled using Trinity

software [16]. The contigs sequences (>1000 bp size) were compared with invertebrate genome data using the Basic Local Alignment Search Tool (BLAST) software from NCBI (https://blast.ncbi.nlm.nih.gov), then assembled via Bowtie2 software [17; 18] with the mitochondrial genome of different species of *Culex* spp. that were retrieved from the GenBank of the National Center of Biotechnology Information (NCBI) nucleotide database. The circular map of the mitochondria was assembled via CHLOROBOX online software (https://chlorobox.mpimp-golm.mpg.de/index.html). The protein-coding genes, transfer RNA genes, ribosomal RNA genes and the non-coding regions were located on the mitochondrial genome sequence operating the MITOS3 online analysis software [19; 20].

Phylogenetic Analysis

The mitochondrial genomes of various species of *Culex* spp. and the Iraqi *Culex* sp. were gained from the GenBank database of the NCBI nucleotide databases. Sequence alignments and filtering of the complete mitochondrial genomes were conducted using CLUSTALW software, and the Neighbour-joining tree method in the Molecular Evolutionary Genetics Analysis platform (MEGA; version 11.0.13). A phylogenetic tree was built with bootstrap replicated evaluation nodes 1000 times [21].

Results and Discussion

The Structure Analysis of the mitochondrial genome

The complete mitochondrial genome of the Iraqi *Culex* sp. is 14,856 bp (Figure 1). The genomic components were organized as in other *Culex* spp. This comprises 34 genes: 13 PCGs, 22 tRNA, and rRNA genes (Table 1). The A + T content was 77.09%, while the G+C content was only 22.1%. All PCGs open with the start codon ATN and end with the stop codon TAA. Four of the 13 PCGs are encoded on the L-strand (nad1, nad4, nad4L, nad5), while the rest are on the H-strand (Figure 1). The MITOS3 analysis (Table 1) showed that the Iraqi common house mosquito is *C. pipiens*. It was deposited in the GenBank database of the NCBI under accession number OQ533057.

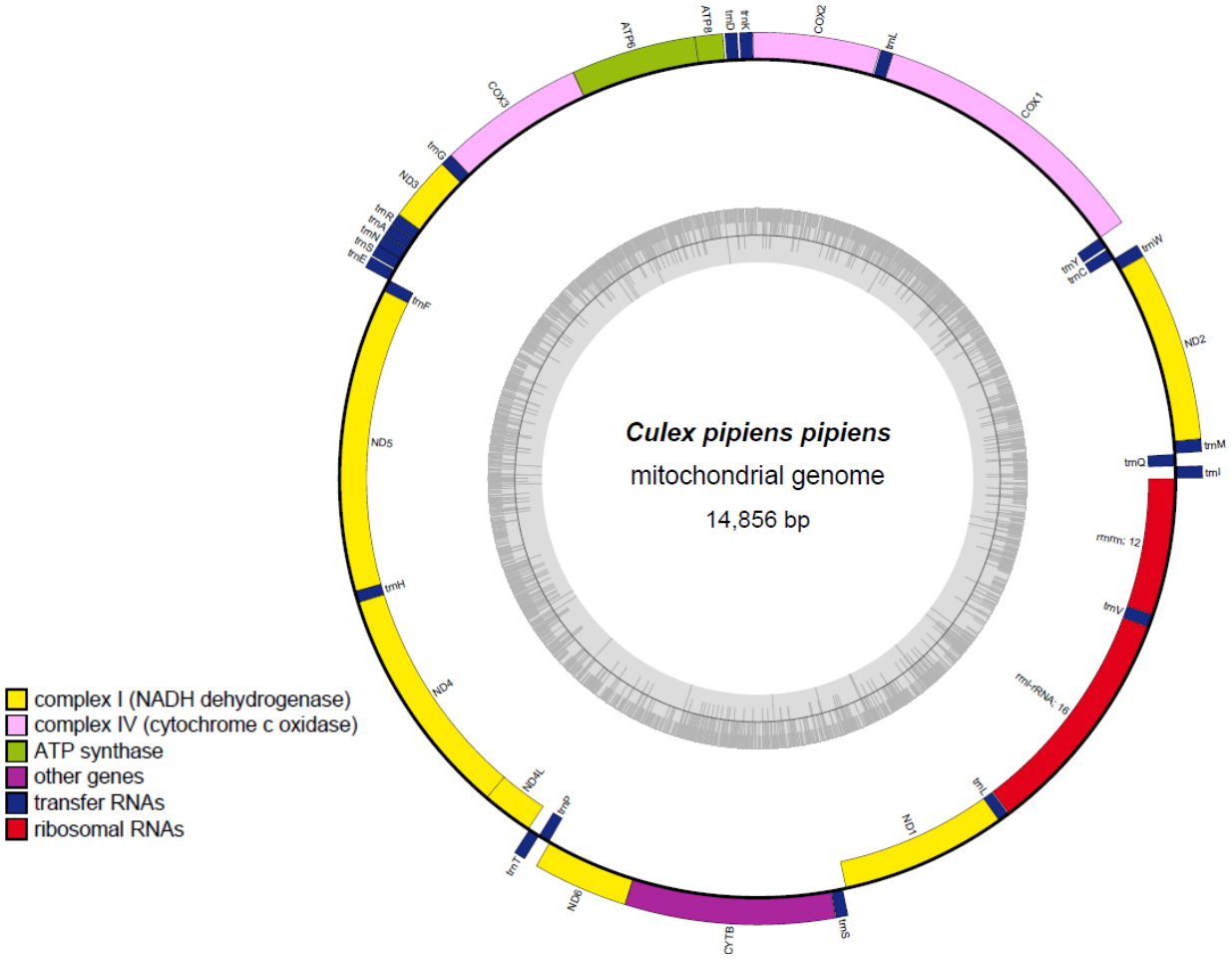

Figure (1): Circular mitogenome diagram of the Iraqi *C. pipiens*. The outside of the circular is the heavy chain, whereas the inside is the light chain. Various colours denote different gene groups: Yellow demonstrates the NADH dehydrogenase genes, pink represents the cytochrome c oxidase genes, green shows the ATP synthase genes, the red displays other genes. The purple represents the transport RNA, while the red represents the two ribosomal RNAs.

Table (1): The mitochondrial genome organization of the Iraqi *C. pipiens*

Name	Start	Stop	Strand	Length
trnI(atc)	1	69	H	69
trnQ(caa)	70	138	L	69
trnM(atg)	142	210	H	69
nad2	232	1209	H	978
trnW(tga)	1235	1303	H	69
trnC(tgc)	1303	1369	L	67
trnY(tac)	1382	1447	L	66
cox1	1452	2960	H	1509
cox2	3055	3735	H	681
trnD(gac)	3821	3888	H	68
atp8	3889	4047	H	159
atp6	4044	4715	H	672
cox3	4724	5509	H	786
nad3	5585	5923	H	339
trnR(cga)	5931	5994	H	64
trnA(gca)	5995	6060	H	66
trnN(aac)	6061	6127	H	67
trnS1(agc)	6130	6196	L	67
trnE(gaa)	6198	6263	H	66
trnF(ttc)	6262	6328	L	67
nad5	6347	8032	L	1686
trnH(cac)	8072	8137	L	66
nad4	8146	9480	L	1335
nad4l	9477	9743	L	267
trnT(aca)	9776	9840	H	65
trnP(cca)	9841	9906	L	66
nad6	9912	10415	H	504
cob	10427	11557	H	1131
trnS2(tca)	11575	11627	H	53
nad1	11655	12587	L	933
trnL1(cta)	12600	12660	L	61
rrnL	12632	13988	L	1357
trnV(gta)	14000	14071	L	72
rrnS	14071	14850	L	780

Genes constructions

The entire length of the 13 PCGs nucleotide sequences in the Iraqi *C. pipiens* is 10,980 bp. Therefore, the start codon of the majority of the PCGs is a typical ATN codon. However, the start codon of the *cox1* gene has not been determined yet. On the

other hand, the stop codons of the 13 PCGs are TAA. Even though most of these stop codons are incomplete (T-) codons, they are frequently discovered in metazoan mito-genomes. This is probably finalized by post-transcriptional polyadenylation [22]. There are 22 tRNA genes that fluctuate from 53 bp (trnS2) to 72 bp (trnV). Of these 22 tRNA genes, 14 are encoded by the H-strand, while only eight are encoded by the L-strand (Figure 1; Table 1). On the other hand, the two rRNA genes are encoded by the L-strand with a size of 780 bp for the rrnS and 1357 bp for the rrnL.

Phylogenetic Analysis

The phylogenetic relationships among 22 species of Culex were conducted and relied on the nucleotide sequences of the complete mitochondrial genome. *Bemisia tabaci* was selected as the outgroup. Two isolates of *C. pipiens pipiens* with the Iraqi isolate clustered together in one clade share a close ancestry (Figure 2) with the highest per cent identity (99.37%; Figure 3). Thus, it is more applicable to classify the Iraqi species as *C. pipiens pipiens.*

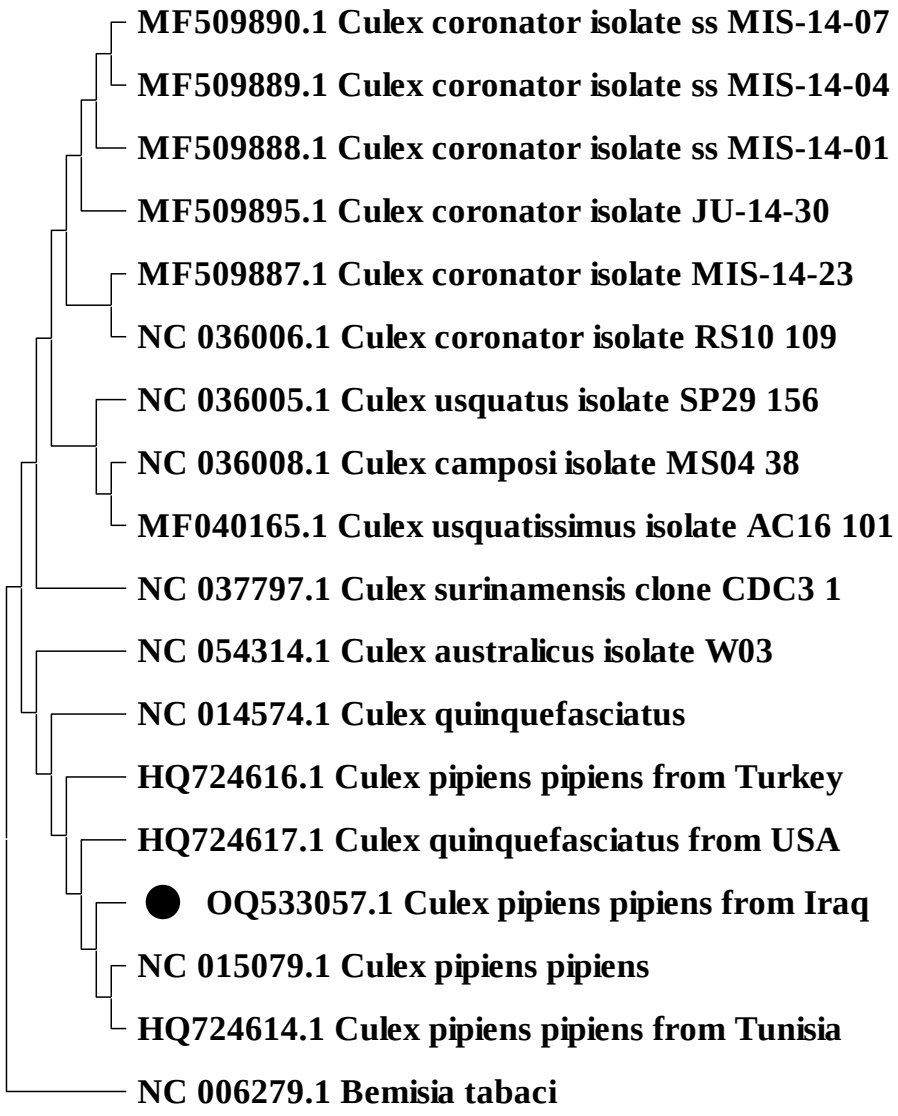

Figure (2): Phylogeny of *C. pipiens*. The phylogenetic tree was built based on the mitogenome's complete genome sequences. *Bemisia tabaci* was the outgroup

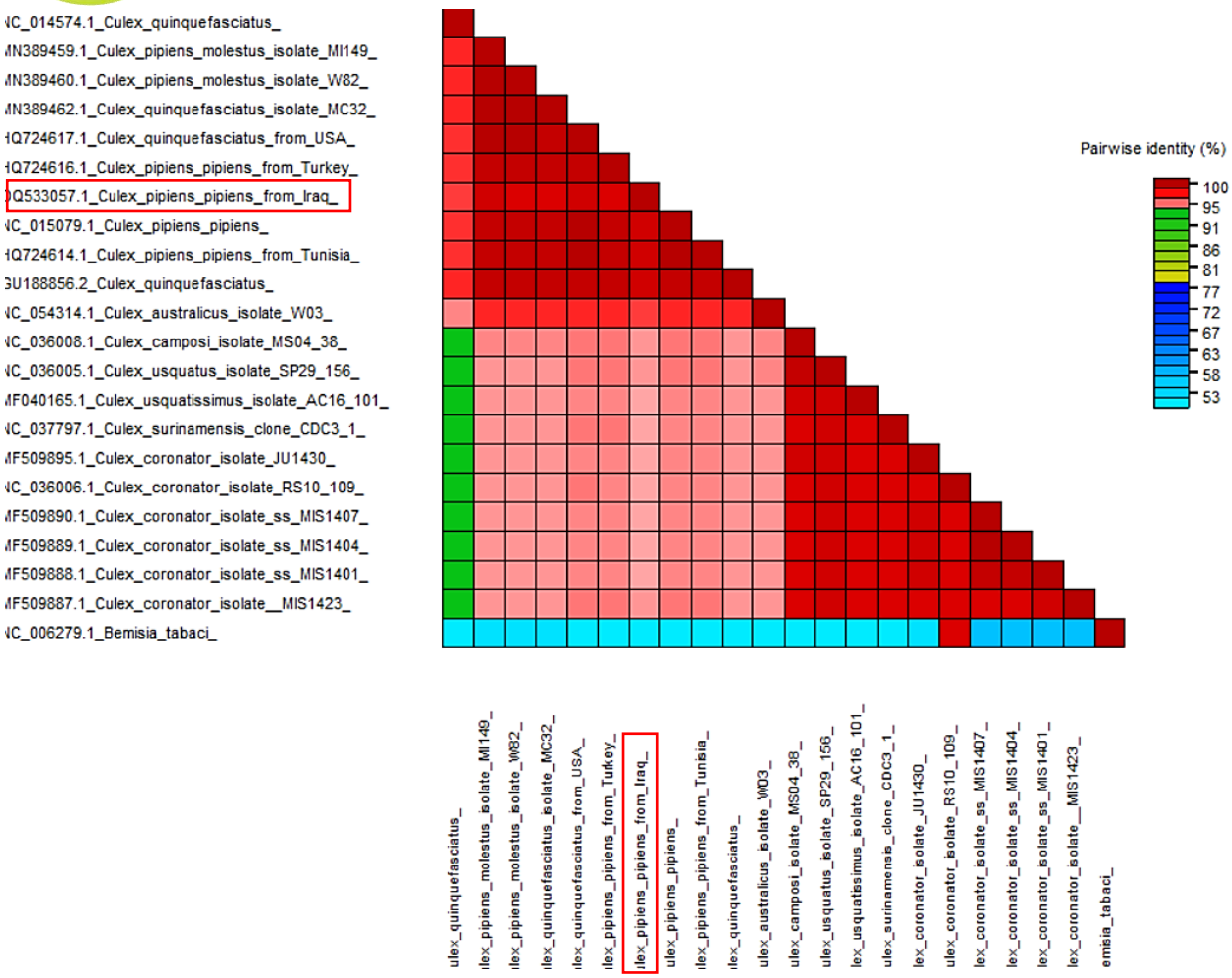

Figure (3): Pairwise identity of *C. pipiens*. The graphical illustration was built based on BLASTn analysis of the Iraqi *C. pipiens* mitogenome sequence (indicated by red border rectangles) compared with other *Culex* species. This graphical diagram was created using Sequence Demarcation Tool version 1.2.

To our knowledge, the first whole mitochondrial genomes of *C. pipiens pipiens* was determined in Iraq and compared with other Culex species. The length of the mitogenome sequence of *C. pipiens* is 14,856 bp. It is a circular structure containing 13 protein-coding genes PCGs, 22 tRNA genes, and two rRNA genes in addition to non-coding regions. All PCGs open with the start codon ATN and end with the stop codon TAA. Comparable to other invertebrate mitogenomes, all PCGs use ATN as the start codon, except for the *cox1* gene (undetermined). In the current work, the phylogenetic analysis indicated clearly that the Iraqi common house mosquito is *C. pipiens pipiens*, belonging to the family Culicidae. The results of this study provide valuable molecular facts for additional considerations concerning the population genetics and evolutionary biology of *C. pipiens pipiens* in Iraq. Thus, further studies are needed.

References

1) Foster, W. A. and Walker, E. D. (2019). Mosquitoes (Culicidae): pp, 261-325. In Mullen, K. G., Durden, L. A. (Eds.). Medical and veterinary entomology. Academic Press, Amsterdam, The Netherlands.
2) Shorb, H. (2013). Fundamentals of medical and veterinary entomology. Academic Library, Cairo, Arab Republic of Egypt.
3) Sulayman, K. A. and Yassn, A. T. (2020). Investigation of Toxicity of some Plant Extracts on Mosquito Larvae *Culex pipiens molestus*. Journal of Al Rafedeen Sciences 29(1):41-49.
4) Al-Ahmad, A. M. N. (2003). Diseases transmitted by mosquitoes. Saudi Society for Agricultural Sciences, Riyadh, Saudi Arabia.
5) Kimura, M.; Darbro, J.M. and Harrington, L.C. (2010). Avian malaria parasites share congeneric mosquito vectors. Journal of Parasitology 96(1):144-51.
6) Ledesma, N. and Harrington, L. (2011). Mosquito vectors of dog heartworm in the United States: Vector status and factors influencing transmission efficiency. Topics in Companion Animal Medicine 26(4):178-85.
7) Rutledge, C. R. (2008). Mosquitoes (Diptera: Culicidae), 2476-2482. In Capinera J. L. (Ed.). Encyclopedia of Entomology. Springer, Dordrecht, The Netherland.
8) Carpenter, S. J. and La Casse, W.J. (1974). Mosquitoes of North America (North of Mexico). University of California Press, Berkeley, California. The USA.
9) Farajollahi, A. and Price, D. C. (2013). A rapid identification guide for larvae of the most common North American container-inhabiting Aedes species of medical importance. Journal of the American Mosquito Control Association, 29(3): 203-221.
10) Atta, A.R.A.; Jabbar, A. S. and Abdulkader, A. A. (2019). Taxonomic study of some species of flower flies (Diptera: Syrphidae) at Basrah Province. Basrah Journal of Agricultural Sciences, 32(2): 169-175.
11) Yang, T.; Xu, G.; Gu, B.; Shi, Y.; Mzuka, H. L. and Shen, H. (2019). The Complete Mitochondrial Genome Sequences of the Philomycus bilineatus (Stylommatophora: Philomycidae) and Phylogenetic Analysis. Genes 10(3): 198-211.
12) Ye, Z-X; Lahuf, A. A.; Salman, M. D.; Zhang, Y. and Li, J-M.(2022). Draft genome sequence of a novel Iflavirus from Leafhoppers (Exitianus capicola) in Iraq. Microbiology Resource Announcement 11 (5):1-2.
13) Fendt, L.; Röck, A.; Zimmermann, B.; Bodner, M.; Thye, T.; Tschentscher, F.; Owusu-Dabo, E.; Göbel, T.M.; Schneider, P.M. and Parson,W. (2012). MtDNA diversity of Ghana: A forensic and phylogeographic view. Forensic Science International: Genetics 6: 244–249.
14) Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics. babraham.ac.uk/projects/fastqc/.

15) Lahuf, A.A. (2021). RNA-Seq data analysis reveals various viral sequences associated with genome of date palm (Phoenix dactylifera). IOP Conference Series: Earth and Environmental Science 735: 012027.

16) Grabherr, M. G.; Haas, B. J.; Yassour, M.; Levin, J. Z.; Thompson, D. A.; Amit, I.; … et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology 29(7): 644-652.

17) Langmead, B. and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods 9(4): 357-359.

18) Abass, M.O. and Lahuf, A.A. (2022). First Report of the Satellite DNA Beta Associated with Tomato Yellow Leaf Curl Virus–Mild on Tomato in Iraq. Plant Health Progress 23 (4): 480-481.

19) Bernt, M.; Donath, A.; Jühling, F.; Externbrink, F.; Florentz, C.; Fritzsch, G.; Pütz, J.; Middendorf, M. and Stadler, P.F. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. Molecular Phylogenetics and Evolution 69: 313–319.

20) Salman, M. D. and Lahuf, A. A. (2022). A New Procedure to identify plant RNA viruses associated with the whitefly (*Bemisia tabaci*) using next-generation sequencing. Arab Journal of Plant Protection 40 (2): 169-174.

21) Kumar, S., Stecher, G., and Tamura, K. J. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

22) Ojala, D.; Montoya, J. and Attardi, G. (1981). tRNA punctuation model of RNA processing in human mitochondrial. Nature, 290: 470–474.

23) Al-Tememe, Z A. M.; Lahuf, A. A.; Abdalmoohsin, R. G. and Al-Amirry, A. T. (2019). Occurrence, identification, pathogenicity and control of *Neoscytalidium dimidiatum* fungus, the causal agent of sooty Canker on eucalyptus camaldulensis in Kerbala province of Iraq. Plant Archives 19 (Supplement 1):31-38.