Available online freely at www.isisn.org

Bioscience Research

Print ISSN: 1811-9506 Online ISSN: 2218-3973

Journal by Innovative Scientific Information & Services Network

RESEARCH ARTICLE BIOSCIENCE RESEARCH, 2019 16(2): 1944-1951. **OPEN ACCESS**

Mitochondrial Genetic Variations of *Musca domestica* in 10 countries

Tanawat Chaiphongpachara.

Department of Public Health and Health Promotion, College of Allied Health Science, Suan Sunandha Rajabhat University, Samut Songkhram 75000, **Thailand**.

***Correspondence:** tanawat.ch@ssru.ac.th **Accepted:** 21 May. 2019 **Published online:** 07 June 2019

The house fly, *Musca domestica* Linnaeus, 1758 is an insect vector of more than 100 human and animal pathogens. In this study, 41 cytochrome c oxidase subunit I (COI) *M. domestica* sequences from 10 countries across four continents, including Canada (North America), USA (North America), China (Asia), Bangladesh (Asia), Saudi Arabia (Asia), India (Asia), Thailand (Asia), South Korea (Asia), Australia (Australia), and Ecuador (South America), were used to estimate genetic variation and structure. There were 23 polymorphic sites, and the average overall values of haplotype diversity and nucleotide diversity were 0.777 ± 0.049 and 0.003 ± 0.003, respectively. The total number of haplotypes was 15, which consisted of two shared and 13 singular haplotypes. The *M. domestica* population genetic structure analysis based on analysis of molecular variance demonstrated 66.13% genetic variation (Φ_{ST} = 0.279), which was significant based on 10,000 permutations ($p < 0.05$). Our results have shed light on genetic differences across *M. domestica* populations around the world, which will yield insight into better vector control methods.

Keywords: ***Musca domestica*, genetic variation, cytochrome c oxidase subunit I**

INTRODUCTION

The house fly *Musca domestica* Linnaeus, 1758 (order Diptera and family Muscidae) (Killick-Kendrick, 1996), originated from the savannahs of Central Asia and then spread globally (Khamesipour et al., 2018). At present, *M. domestica* is one of the most common fly species and is typically associated with humans and found in areas of human activities such as hospitals, food markets, slaughterhouses, food shops or restaurants, and animal farms (Killick-Kendrick, 1996; Khamesipour et al., 2018). In addition, topography also affects *M. domestica* population densities. Recently, medically important fly species were explored in the coastal, urban, upper and lower alluvial, and mountainous areas of central Thailand, with the lower alluvial areas noted to have the highest population of *M. domestica* (Chaiphongpachara et al., 2018).

M. domestica is a vector of more than 100 human and animal pathogens, including bacteria such as *Bacillus* spp., *Coccobacillus* spp., *Staphylococcus* spp., *Micrococcus* spp., *Streptococcus* spp., *Acinetobacter* spp., *Enterobacter* spp., *Proteus* spp., *Escherichia* spp., and *Klebsiella* spp. (Nazni et al., 2005); fungi such as *Aspergillus flavus*, *A. niger* var. niger, *Penicillium corylophilum*, *P. fellutanum*, *Cladosporium cladosporioides*, *Fusarium* spp., *Alternaria alternata*, *Curvularia brachyspora*, *Mycelia sterilia*, and Mucorales order species (De Senna Nunes Sales et al., 2002); viruses such as rotavirus (Tan et al., 1997) and porcine reproductive and respiratory syndrome virus (Otake et al., 2004); protozoans such as *Cryptosporidium parvum* (Graczyk et al., 1999), *Blastocystis* spp., *Giardia intestinalis*, *Endolimax nana*, *Cyclospora cayetanensis*, *Toxocara* spp.,

and *Entamoeba histolytica* (Muñoz and Rodríguez, 2015); and helminths such as *Ascaris lumbricoides, Necator americanus*, and *Fasciola hepatica* (Onyenwe et al., 2016). Given the plethora of pathogens that can be transmitted, gleaning a better understanding of *M. domestica* population dynamics is critical.

Genetic variations of global house fly populations yield important information that will contribute to vector control. The short sequence fragment of the mitochondrial gene cytochrome c oxidase subunit I (COI) has been used in DNA barcoding studies, including species identification (Weeraratne et al., 2018). Currently, standardized COI sequence data of global *M. domestica* populations were collected in DNA reference libraries, including the Barcode of Life Data (BOLD) systems and the National Center for Biotechnology Information (NCBI) GenBank.

Although *M. domestica* is spread worldwide, little is known about its genetic variability and structure. Herein, we investigated *M. domestica* genetic variation across globally distinct populations based on COI sequences from DNA reference libraries. Our results have augmented the current landscape of genetically distinct *M. domestica* populations around the world, which will be critical in providing insight into vector control and surveillance.

MATERIALS AND METHODS

M. domestica COI Sequence Selection

In this study, 41 *M. domestica* COI sequences from 10 countries across four continents were used to estimate and assess genetic variation and structure (Supplemental Table 1).

Table 1. *M. domestica* COI sequences used in this study

Continents	Countries	Databases	Accession no.
North America	Canada	NCBI	HQ982491.1
North America	Canada	NCBI	KM571920.1
North America	Canada	NCBI	KM571474.1
North America	Canada	NCBI	KM570927.1
North America	Canada	NCBI	KM570270.1
North America	Canada	NCBI	HM389240.1
North America	Canada	NCBI	HM389239.1
North America	USA	NCBI	KC617839.1
North America	USA	NCBI	KC617838.1
North America	USA	NCBI	KC617837.1
North America	USA	NCBI	KC617836.1
Asia	China	NCBI	KJ129470.1
Asia	China	NCBI	KJ129467.1
Asia	China	NCBI	KJ129464.1
Asia	China	NCBI	KJ129463.1
Asia	China	NCBI	KJ129461.1
Asia	Bangladesh	NCBI	MG572239.1
Asia	Bangladesh	NCBI	MG557665.1
Asia	Saudi Arabia	NCBI	KU578307.1
Asia	Saudi Arabia	NCBI	KU578305.1
Asia	Saudi Arabia	NCBI	KU578306.1
Asia	Saudi Arabia	NCBI	KU578302.1
Asia	Saudi Arabia	NCBI	KU578301.1
Asia	Saudi Arabia	NCBI	KU578300.1
Asia	India	NCBI	KC427132.1
Asia	India	BOLD	AGIRI228-17
Asia	Thailand	BOLD	ENTJR422-08
Asia	Thailand	BOLD	ENTJR423-08
Asia	Thailand	BOLD	ENTJR424-08
Asia	Thailand	BOLD	ENTJR425-08
Asia	Thailand	BOLD	ENTJR430-08
Asia	South Korea	NCBI	JX861432.1
Asia	South Korea	NCBI	JX861433.1
Asia	South Korea	BOLD	GBDP15322-14

Australia	Australia	BOLD	DIQTB495-12
Australia	Australia	BOLD	DIQTB581-12
Australia	Australia	BOLD	DIQTB602-12
South America	Ecuador	BOLD	FBCTW047-14
South America	Ecuador	BOLD	FBCTW048-14
South America	Ecuador	BOLD	FBCTW064-14
South America	Ecuador	BOLD	FBCTW227-14

All selected COI sequences were from the NCBI GenBank and BOLD DNA reference libraries. The Basic Local Alignment Search Tool (BLAST) was used to confirm species identity of the COI sequences, and if sequences did not match with *M. domestica*, those samples were removed from our analysis.

Statistical Analysis

All *M. domestica* COI sequences were aligned using Clustal X and manually edited. The genetic variation of *M. domestica* populations from each country was evaluated to determine the numbers of haplotypes (N), polymorphic sites, nucleotide (π) diversity, and haplotype (h) diversity using the DnaSP6 software (Rozas et al., 2017).

Analysis of molecular variance (AMOVA) based on 10,000 permutations as significance at p-values < 0.05 was used to evaluate population genetic structure, including the percentage of sequence divergence within and between 10 *M. domestica* populations using the Arlequin 3.5.2.2 software (Excoffier and Lischer, 2010). *M. domestica* populations were grouped by continents. Pairwise F_{ST} values based on 10,000 permutations as significance at p-values < 0.05 using Arlequin 3.5.2.2 (Excoffier and Lischer, 2010) were used to assess short-term genetic distances between *M. domestica* populations.

The maximum likelihood method with 10,000 bootstrap replicates was used for phylogenetic analysis of the 41 *M. domestica* COI sequences and was calculated using the MEGA7 software (Kumar et al., 2016). In addition, the median-joining haplotype network (Bandelt et al., 1999) was built to assess geographic relationships based on statistical parsimony using Arlequin 3.5.2.2 (Excoffier and Lischer, 2010) and the Network 5.0.1.1 software available at http://www.fluxus-engineering.com.

RESULTS

Genetic Diversity

The COI sequences (639 bp) of 41 *M. domestica* samples were obtained from NCBI GenBank and BOLD and then analyzed. The sequence variants (haplotypes) are shown in Figures 2 and 3. There were 23 polymorphic sites, and the average overall values of haplotype diversity and nucleotide diversity were 0.777 ± 0.049 and 0.003 ± 0.003, respectively (Table 2). The total number of haplotypes was 15. There were two shared haplotypes, H5 (shared with Canada, USA, China, Saudi Arabia, and India) and H6 (shared with Canada, China, Bangladesh, Saudi Arabia, Thailand, South Korea, Australia, and Ecuador), and 13 singular haplotypes, including H1 (specific to Ecuador), H2 (specific to Saudi Arabia), H3 (specific to Australia), H4 (specific to Australia), H7 (specific to Bangladesh), H8 (specific to China), H9 (specific to China), H10 (specific to South Korea), H11 (specific to Ecuador), H12 (specific to Saudi Arabia), H13 (specific to Ecuador), H14 (specific to India), and H15 (specific to China) (Table 3).

Population Genetic Structure

M. domestica genetic structure was analyzed using 10 geographic populations, Canada, USA, China, Bangladesh, Saudi Arabia, India, Thailand, South Korea, Australia, and Ecuador. AMOVA indicated 66.13% of genetic variation (Φ_{ST} = 0.279), which was statistically significant ($p < 0.05$) (Table 4). Pairwise F_{ST} analysis showed significant genetic differences between the populations of China and USA, Saudi Arabia and USA, Thailand and Canada, Thailand and USA, Thailand and China, Thailand and India, South Korea and Canada, South Korea and USA, Australia and Canada, Australia and USA, Ecuador and Canada, Ecuador and USA, and Ecuador and Thailand ($p < 0.05$) (Table 5).

Phylogenetic Tree and Median-Joining Haplotype Network

Figure 1 is the phylogenetic tree analysis based on the maximum likelihood method with 10,000 bootstrap replicates, which did not show a genetic relationship between countries, whereas Figure 2 is the median-joining haplotype network tree, which showed a relationship of haplotypes divided into two clusters.

Table 2. ***M. domestica*** **haplotypes and nucleotide diversity**

Countries	No	No. haplotypes	No. polymorphic sites	Haplotype diversity (h) (mean ± SD)	Nucleotide diversity (π) (mean ± SD)
Canada	7	2	1	0.286±0.196	0.000±0.000
USA	4	1	0	0.000±0.000	0.000±0.000
China	5	5	6	1.000±0.126	0.004±0.003
Bangladesh	2	2	4	1.000±0.500	0.006±0.005
Saudi Arabia	6	4	5	0.800±0.172	0.003±0.002
India	2	2	3	1.000±0.500	0.005±0.004
Thailand	5	1	0	0.000±0.000	0.000±0.000
South Korea	3	2	4	0.667±0.134	0.004±0.003
Australia	3	3	6	1.000±0.272	0.006±0.004
Ecuador	4	4	11	1.000±0.177	0.009±0.005
Total	41	15	23	0.777±0.049	0.003±0.003

Table 3. ***M. domestica*** **haplotype distributions from 10 countries**

Haplotypes	Countries										Total
	CAN	USA	CHI	BAN	SAU	IND	THA	KOR	AUS	ECU	
H1	-	-	-	-	-	-	-	-	-	1	1
H2	-	-	-	-	1	-	-	-	-	-	1
H3	-	-	-	-	-	-	-	-	1	-	1
H4	-	-	-	-	-	-	-	-	1	-	1
H5	6	4	1	-	1	1	-	-	-	-	13
H6	1	-	1	1	3	-	5	2	1	1	15
H7	-	-	-	1	-	-	-	-	-	-	1
H8	-	-	1	-	-	-	-	-	-	-	1
H9	-	-	1	-	-	-	-	-	-	-	1
H10	-	-	-	-	-	-	-	1	-	-	1
H11	-	-	-	-	-	-	-	-	-	1	1
H12	-	-	-	-	1	-	-	-	-	-	1
H13	-	-	-	-	-	-	-	-	-	1	1
H14	-	-	-	-	-	1	-	-	-	-	1
H15	-	-	1	-	-	-	-	-	-	-	1
Total	7	4	5	2	6	2	5	3	3	4	41

Table 4. AMOVA of ***M. domestica*** **from 10 geographical populations**

Source of variation	df	Sum of squares	Variance components	Percentage of variation	Φ	p-value
Among groups	2	3.770	0.123Va	27.95	Φ_{CT} = 0.082	0.000*
Among populations	7	2.742	0.026Vb	5.92	Φ_{SC} = 0.339	0.192
Within populations	31	9.024	0.291Vc	66.13	Φ_{ST} = 0.279	0.019*
Total	40	15.537	0.440			

Table 5. Pairwise comparisons of *M. domestica* genetic diversity among 13 countries

	CAN	USA	CHI	BAN	SAU	IND	THA	KOR	AUS	ECU
CAN	-									
USA	-0.098 (0.100)	-								
CHI	0.239 (0.068)	0.322 (0.050)*	-							
BAN	0.527 (0.086)	0.724 (0.069)	-0.111 (0.100)	-						
SAU	0.324 (0.051)	0.442 (0.033)*	-0.034 (0.745)	-0.154 (0.100)	-					
IND	0.161 (0.405)	0.384 (0.335)	-0.111 (0.100)	0.000 (0.100)	0.062 (0.459)	-				
THA	0.804 (0.013)*	1.000 (0.006)*	0.375 (0.045)*	0.474 (0.282)	0.161 (0.179)	0.773 (0.045)*	-			
KOR	0.552 (0.035)*	0.724 (0.028)*	0.008 (0.464)	-0.200 (0.100)	-0.116 (0.100)	0.207 (0.397)	0.189 (0.393)	-		
AUS	0.454 (0.031)*	0.579 (0.030)*	-0.071 (0.100)	-0.200 (0.100)	-0.059 (0.100)	0.000 (0.100)	0.423 (0.110)	-0.071 (0.100)	-	
ECU	0.410 (0.015)*	0.500 (0.029)*	-0.053 (0.100)	-0.143 (0.100)	-0.018 (0.572)	0.000 (0.100)	0.394 (0.048)*	-0.020 (0.482)	-0.091 (0.100)	-

* = significant differentiation ($p < 0.05$)

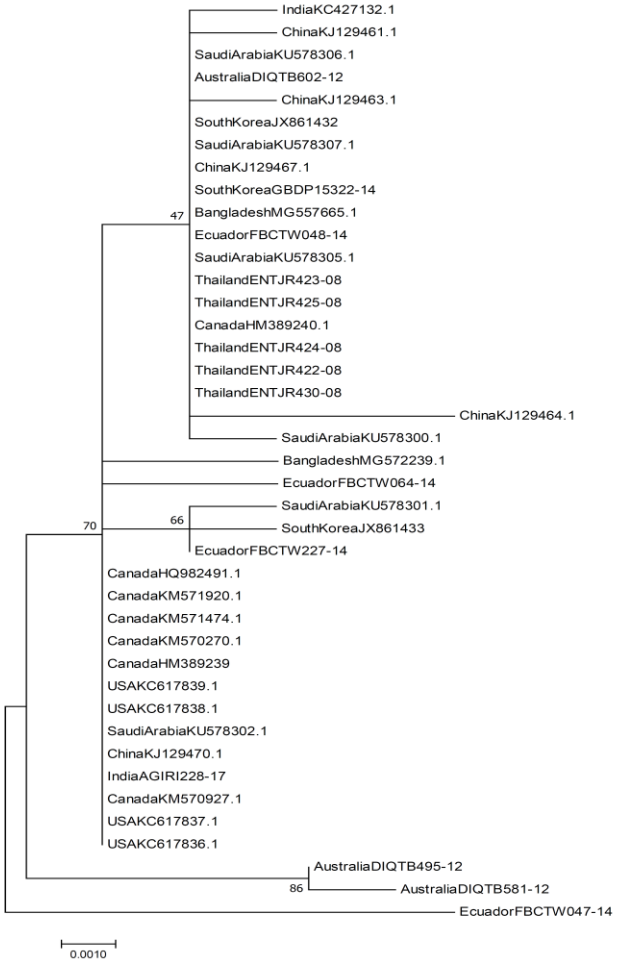

Figure 1. Phylogenetic tree of 41 COI sequences of *M. domestica* from 10 countries based on the maximum likelihood method

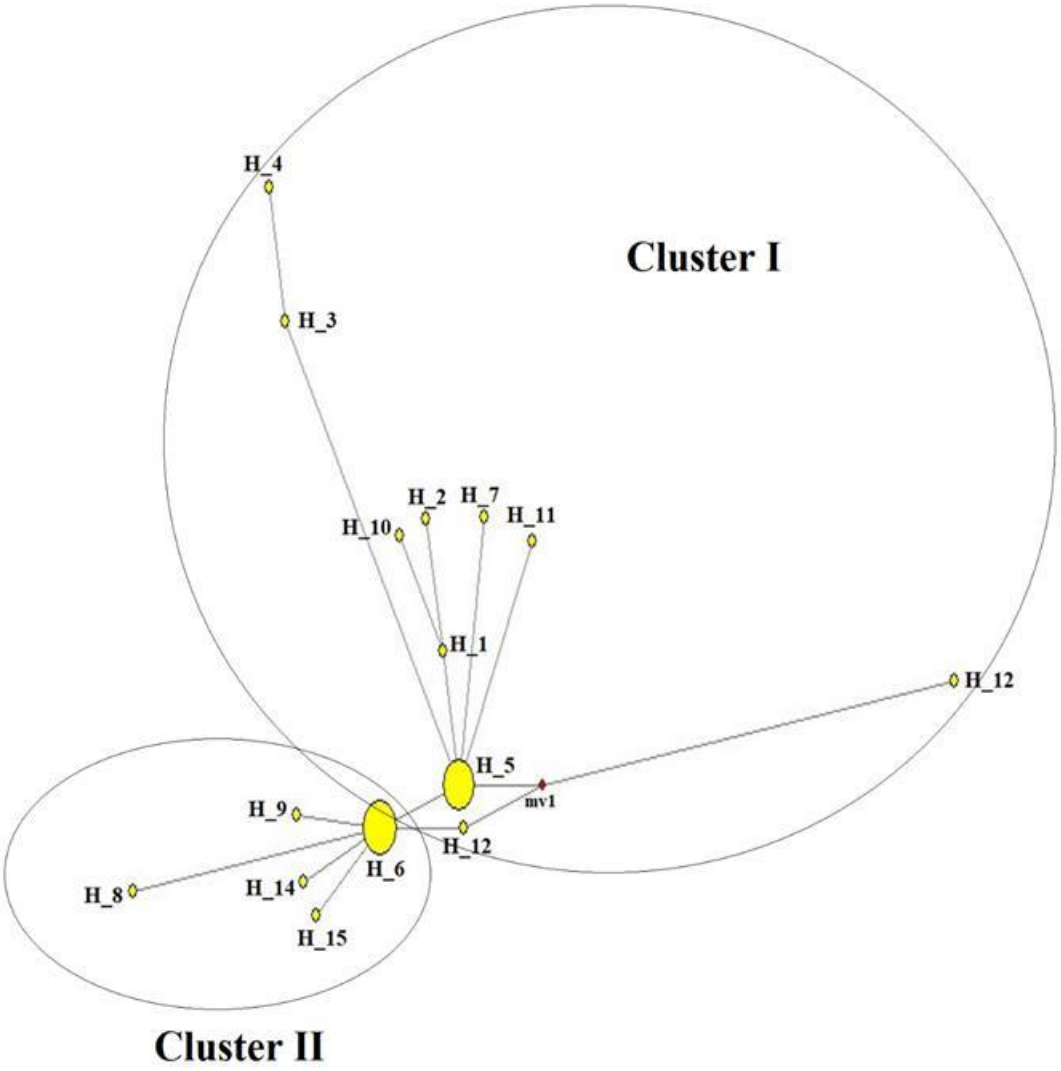

Figure 2. A minimum spanning network based on 15 haplotypes of 41 COI sequences of *M. domestica* from 10 countries, each circle (yellow color) representing a haplotype (the size of each circle proportional to the frequency of each haplotype and branch lengths proportional to the number of nucleotide changes between haplotypes)

Cluster I included 10 haplotypes with H5 as a central haplotype, and cluster II included five haplotypes with H6 as a central haplotype.

DISCUSSION

Over the last decade, DNA barcoding has become a key method employed for animal species identification and genetic variation in which a short DNA sequence is used for analysis (Yang et al., 2018). Sequences are collected and deposited into publicly available databases which useful for future studies of species populations (Lim et al., 2009). In this study, genetic diversity was analyzed using 41 COI sequences from 10 *M. domestica* populations around the world, which were retrieved from the NCBI GenBank and BOLD databases.

We identified 15 haplotypes (consisting of 13 singular haplotypes and two shared haplotypes), and the overall average of haplotype and nucleotide diversities was reportedly low in Canada, USA, and Thailand. Previous reports indicated that low diversity of *M. domestica* is affected by overwintering population collapse in temperate regions such as Canada and USA (Marquez and Krafsur, 2002). Meanwhile, our

finding showed that in Thailand, which is located in a tropical region, this low diversity may be due to the low number of *M. domestica* samples.

There was 66.13% of genetic variation identified, with genetic differences between populations in China and USA, Saudi Arabia and USA, Thailand and Canada, Thailand and USA, Thailand and China, Thailand and India, South Korea and Canada, South Korea and USA, Australia and Canada, Australia and USA, Ecuador and Canada, Ecuador and USA, and Ecuador and Thailand. The differences in genetic structure of *M. domestica* populations could be due to the adaptation of the vector to various environments and habitats. Our results are in accordance with Cummings and Krafsur (2005), which found pairwise differentiation based on zoogeographical regions of *M. domestica* populations, particularly from Afrotropical, Indo-Malayan, Nearctic, Neotropical, and Palearctic regions.

The median-joining haplotype network tree revealed two clusters: Cluster I with H5 as a central haplotype and Cluster II with H6 as a central haplotype. This result illustrated that most haplotypes were derived from the central haplotypes (H5 in Cluster I and H6 in Cluster II), which may be the consequence of a single mutational change in *M. domestica*.

CONCLUSION

This study demonstrated mitochondrial genetic variations across *M. domestica* populations around the world. Compared to other geographical regions, Nearctic *M. domestica* populations possessed lower genetic diversity values. In addition, genetic structure was significantly distinct within *M. domestica* populations and among groups from different regions. These results enhance our understanding of *M. domestica* genetics, which can be used to improve vector surveillance and control efforts.

CONFLICT OF INTEREST

The author declared that present study was performed in absence of any conflict of interest.

ACKNOWLEGEMENT

The author thanks the College of Allied Health Sciences, Suan Sunandha Rajabhat University, Thailand, for its support.

AUTHOR CONTRIBUTIONS

TC designed, performed data analysis and wrote the manuscript.

REFERENCES

Bandelt HJ, Forster P, Röhl A, 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16(1): 37-48.

Chaiphongpachara T, Laojun S, Jongvisuttisan N, Tubsamut P, Dasom A, 2018. A checklist of medically important flies (Order : Diptera) in the central region of Thailand. Biodiversitas 19(6): 2134–2139.

Cummings MA, Krafsur ES, 2005. Spatial diversity in mitochondrial cytochrome c oxidase in house flies. Med Vet Entomol 19(1): 53-9.

De Senna Nunes Sales M, Da Costa GL, Pinheiro Bittencourt VRE, 2002. Isolation of fungi in *Musca domestica* linnaeus, 1758 (Diptera: Muscidae) captured at two natural breeding grounds in the municipality of Seropédica, Rio de Janeiro, Brazil. Mem Inst Oswaldo Cruz 97(8): 1107-1110,

Excoffier L, Lischer HEL, 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10(3): 564-567.

Graczyk TK, Cranfield MR, Fayer R, Bixler H, 1999. House flies (*Musca domestica*) as transport hosts of cryptosporidium parvum. Am J Trop Med Hyg 61(3): 500-504.

Khamesipour F, Lankarani KB, Honarvar B, Kwenti TE, 2018. A systematic review of human pathogens carried by the housefly (*Musca domestica* L.). BMC Public Health 18(1): 1049.

Killick-Kendrick R, 1996. Medical entomology for students. Trans R Soc Trop Med Hyg 90 (5): 590 pp.

Kumar S, Stecher G, Tamura K, 2016. MEGA7: Molecular Evolutionary Genetics

Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33(7): 1870-1874.

Lim J, Kim SY, Kim S, Eo HS, Kim CB, Paek WK, Kim W, Bhak J, 2009. BioBarcode: A general

DNA barcoding database and server platform for Asian biodiversity resources. BMC Genomics 10 (3): S8.

Marquez JG, Krafsur ES, 2002. Gene flow among geographically diverse housefly populations (*Musca domestica* L.): a worldwide survey of mitochondrial diversity. J. Hered 93(4): 254-9.

Muñoz DJ, Rodríguez R, 2015. Bacterial and parasite agents in adult housefly *Musca domestica* collected in El Peñón, Sucre State, Venezuela. Rev Cient Fac Ciencias Vet Univ del Zulia 25(2): 159-166.

Nazni WA, Seleena B, Lee HL, Jeffery J, T Rogayah TA, Sofian MA, 2005. Bacteria fauna from the house fly, *Musca domestica* (L.). Trop Biomed 22(2): 225-31.

Onyenwe E, Okore OO, Ubiaru PC, Abel C, 2016. Housefly-borne helminth parasites of Mouau and its public health implication for the university community. Anim Res Int 13(1): 2352-2358.

Otake S, Dee SA, Moon RD, Rossow KD, Trincado C, Pijoan C, 2004. Studies on the carriage and transmission of porcine reproductive and respiratory syndrome virus by individual houseflies (*Musca domestica*) Vet Rec 154(3): 80-5.

Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-

Onsins SE, Sanchez-Gracia A, 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol 34(12): 3299-3302.

Tan SW, Yap KL, Lee HL, 1997. Mechanical Transport of Rotavirus by the Legs and Wings of Musca domestica (Diptera: Muscidae). J Med Entomol 34(5): 527-31.

Weeraratne TC, Surendran SN, Karunaratne SHPP, 2018. DNA barcoding of morphologically characterized mosquitoes belonging to the subfamily Culicinae from Sri Lanka. Parasit Vectors 11(1): 266.

Yang F, Ding F, Chen H, He M, Zhu S, Ma X, Jiang L, Li H, 2018. DNA Barcoding for the Identification and Authentication of Animal Species in Traditional Medicine. Evidence-Based Complement. Altern Med. Article ID 5160254: 18pp.