Benha Journal of Applied Sciences (BJAS) **print: ISSN 2356–9751**
Vol.(5) Issue(1) Part (1) (2020), (19-24) **online: ISSN 2356–976x**
http://bjas.journals.ekb.eg

Genome Size Variation for some forensically Important Sarcophagidae in Egypt

A.A.El-Hefnawy[1,3], F.F.Abul Dahab[1], A.A.Ibrahim[1], E.M.Salama[1], Sh.H.Mahmoud[2], J.S.Johnston [3] and A.M. Tarone[3]
[1]Entomology Dept., Faculty of Science, Benha Univ., Benha, Egypt
[2]Zoology Dept., Faculty of Science, Menuofia Univ., Shebin Elkom, Egypt.
[3]EntomologyDept., College of Agriculture & Life Sciences, Texas A&M Univ., College Station, TX, United States of America.
E-Mail: ahmed.alhefnawi@fsc.bu.edu.eg

Abstract

Genome size were estimated for three species of forensically relevant flesh flies (Diptera : Sarcophagidae) collected from three different regions in Egypt. Genome sizes of Wohlfahrtia magnifica ranged from 1040.56 in Edfo to 1229 Mbp in Aswan, the Sarcophaga argyrostoma genome was 863 Mbp in Benha to 1045.87 Mbp in Edfo and finally the genome size of Sarcophaga dux was 1552.17 Mbp as estimated for one individual.These estimates are useful not only as preliminary information for some molecular and genomic studies, but also for determination of the species through immatures. Genome size represents an additional tool for species identification in the immature stages, where they are morphologically difficult or impossible to identify. Sarcophagidae can be differentiated by using Cytochrome Oxidase I, but flow cytometry can be faster than sequencing. Genome sizes appear to vary in a discontinuous fashion. This is the first study to produce genome size estimates for Egyptian Sarcophagidae species. More work will need to be done to determine if there is sexual dimorphism, cryptic species, or polyploidy in Wohlfahrtia magnifica and Sarcophaga argyrostoma populations.

Keywords: Forensic entomology, Sarcophagidae, Genome size, Flow cytometry, Egypt.

1. Introduction

Genome sizes have been determined for a large number of species around the world, with ≈ ٥,٠٠٠ animal species completed [7]. Propidium iodide-based flow cytometry (FCM) is becoming a more prevalent method for determining this value. FCM permits measurement of genome sizes for a large number of nuclei, producing an average fluorescence measure for nuclei at all ploidy levels. FCM also allows for a more rapid identification of genome size, which makes it amenable when examining large numbers of individuals.

Flies have been studied extensively for ecological [2] and forensic [1,6] purposes. Most notably calliphorids and sarcophagids have been useful to forensic entomology, that uses of insects in criminal investigations [6]. Insects are often used to estimate minimum postmortem interval (mPMI), given certain assumptions [20], but can also provide additional information, such as geographic origin [14], and postmortem relocation [17]. The primary colonizing insects associated with corpses are predominantly flies [5]; especially flies belonging to the family Calliphoridae and Sarcophagidae.

To accurately estimate the mPMI, species identification is necessary, because the rate of development on remains differs across forensically informative species [6]. In some cases, morphological identification of larvae is nearly impossible, and DNA methods have been employed to aid this process [18, 21, 22 and 23]. This works for many unknown specimens collected as larvae and pupae from different regions in Egypt since sequence-based autapomorphies exist to distinguish between closely related species using DNA sequence analysis. However, sequencing can be time consuming and costly.

The purpose of this study was to demonstrate the utility, reliability, and ease of use of flow cytometry for genome size determinations in taxonomically diverse samples of regional Diptera of forensic importance, and to explore the potential value of such genome value determinations in forensics.

2. Materials and methods

2.1 Insect samples

Unknown samples were collected throughout Benha (Qualiobya Governorate), Edfo and Aswan (Aswan Governorate), Egypt as larvae and pupae (Per [13], we acknowledge that pupation happens as a subset of intra-puparial development and for simplicity sake we refer to pupae when evaluating any intra-puparial individuals). The collected samples were shipped alive to Texas A&M University, Texas, USA and immediately placed in a -80 ^{0}C freezer. A total of 60 specimens were used for this study. Due to the high similarities between the larvae of the Sarcophagidae samples or between the pupal samples, we sequenced Cytochrome oxidase I from about 10 random samples among of these 60 samples, which represented clear morphological types in the samples. Genome sizes were estimated for all samples, whether they were associated with a sequence or not.

2.2 Flow cytometry

The anterior 1/3 was placed in 1 ml of cold Galbraith buffer **[4]** with the head from a female *Drosophila virilis* standard in a 2-ml tissue grinder and stroked 15 times with an A pestle (WRR, Radnor, PA). The solution then was passed through a nylon filter (20 *u*m) and the recovered flow-through was placed in a 1.5-ml microcentrifuge tube on ice. Propidium iodide was added to the solution at a final concentration 25

*u*g/ml. Nuclei were then stained at 4 ^{0}C in the dark for 30 – 40 min.

The mean fluorescence of each sample was determined using CytExpert software version 1.2.8.0 supplied with the CytoFlex flow cytometer set to excite at 432 nm and the nDNA amount in the sample calculated as the ratio of fluorescence in the sample and standard times the amount of nDNA in the standard.

DNA was estimated as mean channel number of the 2C sample divided by the mean channel number of the 2C standard, and this ratio multiplied times the gametic DNA amount in the standard (1C = 328 Mb for *D. virilis*).

2.3 Cytochrome Oxidase I (COI) Sequencing

Genomic DNA was extracted from the second half of the larval or pupal body using a lithium chloride DNA extraction method. They were individually homogenized in 200 *u*l of Buffer A (100 mM Tris-HCl,100 mM EDTA, and 100 mM NaCl) using a pestle and incubated at 65 ^{0}C for 30 min. Following incubation, 400 *u*l of lithium chloride solution (1M KAc and 3M LiCl) was added and samples cooled on ice for 15 min.

Samples were spun at 12,000 rpm for 15 min and the supernatant transferred to 300 *u*l of isopropanol. Samples were mixed and spun for 15 min at 12,000 rpm. Supernatant was removed and pellets washed with 80% ethanol. Pellets were air-dried and resuspended in 150 *u*l of Tris-EDTAbuffer.

PCR conditions were optimized to amplify a region of the cytochrome oxidase I gene (COI) of the flesh fly mitochondrial genome. Nested PCR was used to increase the yield of the products. Insect primers were used in PCR reaction to amplify ≈ 650bp region of the COI gene [12]. These primers were designed using OligoExplorer (Gene-Link).

Forward Primer

C1-J-2183 (5'- CAA CAT TTA TTT TGA TTT TTT GG -3').

Reverse Primer

TL2-N-3014 (5'- CCA ATG CAC TAA TCT GCC ATA TTA -3').

PCR cycling conditions

PCR was performed by 39 cycles of 95 °C for 1 minute, 45 °C for 1 minute and 72 °C for 2 minutes. A final extension at 72 °C for 5 minutes and the reaction was finally stored at 4 ºC. Reactions were performed on a Veriti PCR Thermocycler (Applied Biosystems). A low annealing temperature was used due to the high adenine and thymine content in the insect mitochondrial genome [12].

Sequencing reactions were performed by the Laboratory of Genome Technology at Texas A&M University on an ABI 3130xl Genetic Analyzer. Sequence chromatograms were edited in Finch (http://geospiza.com), exported into MEGA7 for sequence trimming and aligned using Clustal W2 (http://ebi.ac.uk/Tools/msa/clustalw2).

A dataset of Sarcophagidae were assembled and submitted to the NCBI nucleotide database. The consensus sequences have been submitted to GenBank (Accession numbers MN241038- MN241048).

Phylogenetics

An initial phylogenetic tree (phylogram) was obtained by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (3 categories (+G, parameter = 200.0000)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 14 nucleotide sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. Evolutionary analyses were conducted in MEGA7 **[19]** with 1,000 replicates**.**

3. Results

3.1 Genetic Data

DNA was successfully extracted, the COI gene of the mitochondrial DNA was sequenced from ten samples collected from three different region in Egypt (Aswan, Edfo and Benha), and all consensus sequences were uploaded to GenBank (Accession numbers MN241038- MN241048) as shown in Table (1) with their genome size. Using a neighbor-joining bootstrap tree under the Tamura-Nei model **[19]** with 1,000 replicates, all sequences were assigned species identity as shown in Fig (1).

Table (1) Genome size determinations from collected locations of specimens with their identified Accession numbers on GeneBank.

Species	Stage	1C Mb	Accession numbers	Region collection
Sarcophaga dux	**Larvae**	1552.2	MN241039	Aswan
Wohlfahrtia		1040.56	MN241041	Edfo
magnifica		1109.39	MN241043	
		1122.38	MN241045	
		1189.66	MN241048	

Table (1) Continue

Sarcophaga argyrostoma	**Pupa**	1229	MN241038	Edfo
		987.83	MN241040	
		1045.87	MN241042	
		1036	MN241044	
		863	MN241047	Benha

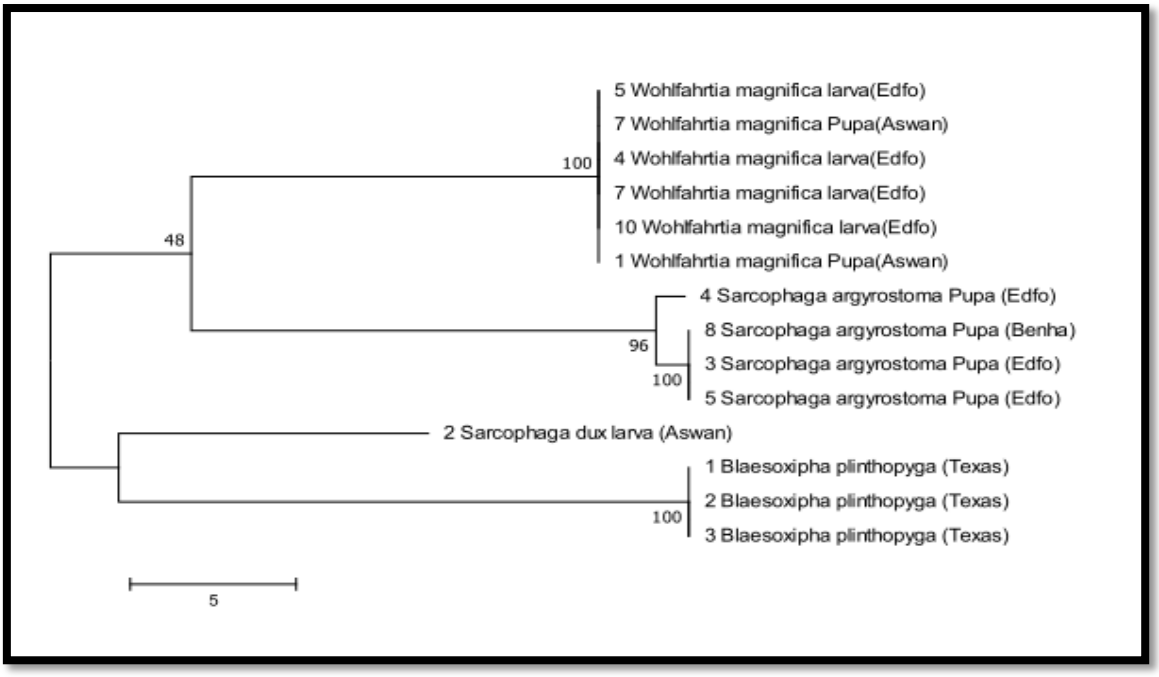

Fig (1) Genetic confirmation using COI. Neighbor-joining bootstrap phylogeny of Sarcophagid flies. The analysis is presented in the form of a phylogram in which branch length indicates the number of base substitutions. The phylogram was based on ≈ 650 bp of the COI gene. Numbers indicate the bootstrap percentage support for each branch (1,000 replicates).

3.1 Genome Size

Larvae and pupae often produced poor results where the histograms have too many peaks and where all except the standard peaks are very small. However there are some that are very good and leave no question about the position of the unknown 2C peak (See Fig. 2). These latter fall into three main categories according to the species identified by COI: First of that is the *W. magnifica* which was associated with ~1041-1122 Mb and ~1190-1229 Mb sizes (may be two sexes) (Figure 2 A&C). The second one is the *S. argyrostoma* was associated with ~863/988 Mb and ~1036-1046 Mb specimens (may be two sexes) (Figure 2 B). The last one is *S. dux* that was flagged as not trustworthy at 1552, but also there are supported peaks around 1500-1770 Mb in the total data set that may justify this value Fig (2), Fig (3) D&E).

From the ten sequenced samples, the genome sizes of *Wohlfahrtia magnifica* as estimated for four individuals were 1C = 1138.2 ± 73.33 (Table 2) ranged from 1040.56 in Edfo to ١٢٢٩ Mbp in Aswan Table (1). The genome size of four individuals of *Sarcophaga argyrostoma* was 1C = 983.18 ± 84.03 Mbp (Table 1) and varied from 863 Mbp in Benha to 1045.87 Mbp in Edfo (Table 1).

The *Sarcophaga dux* genome was 1C= 1552.2 Mbp (lableled as untrustworthy (Fig. 3D), but also associated with similar trustworthy genome sizes in the larger sampling of genome sizes Fig (2).

Table (2) Flow cytometry genome size estimates for three Sarcophagidae species

Species	n	Mean 1C genome size (Mb)	SE (Mb)
Wohlfahrtia magnifica	5	1138.2	73.33
Sarcophaga argyrostoma	4	983.18	84.03
Sarcophaga dux	1	1552.2	0

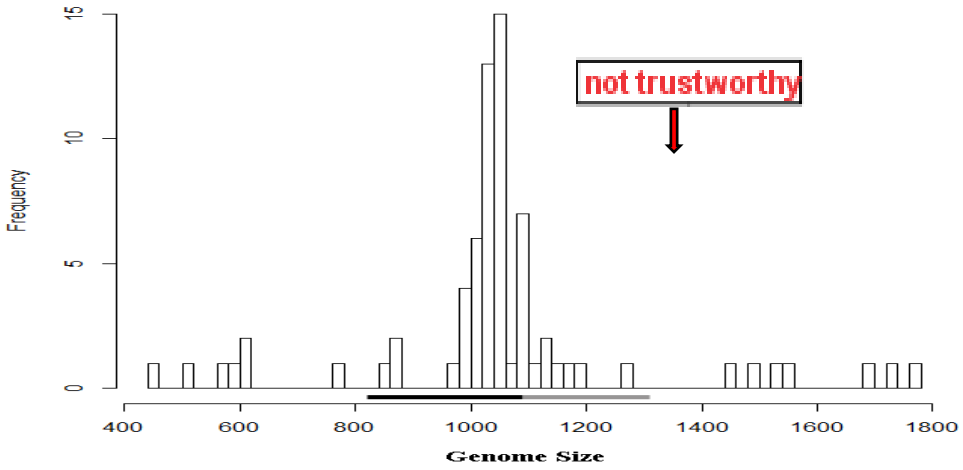

Fig (2) Histogram showing the range of genome sizes run for selected Egyptian Sarcophagidae. Three species were identified by COI sequencing, associated with the values in Table (1). Additional specimens also run and reliable results plus the one sequence result showed similar values as what was produced for sequenced specimens. (The estimated *W. magnifica* genome size range is demonstrated by the gray bar, *S. argyrostoma* by the black bar). The *S. dux* specimen with associated sequence yielded a genome size that should be approached with caution (labeled as not trustworthy), though it should be noted that other samples were found with similar genome sizes, which were not filtered out due to low quality cytometry results.

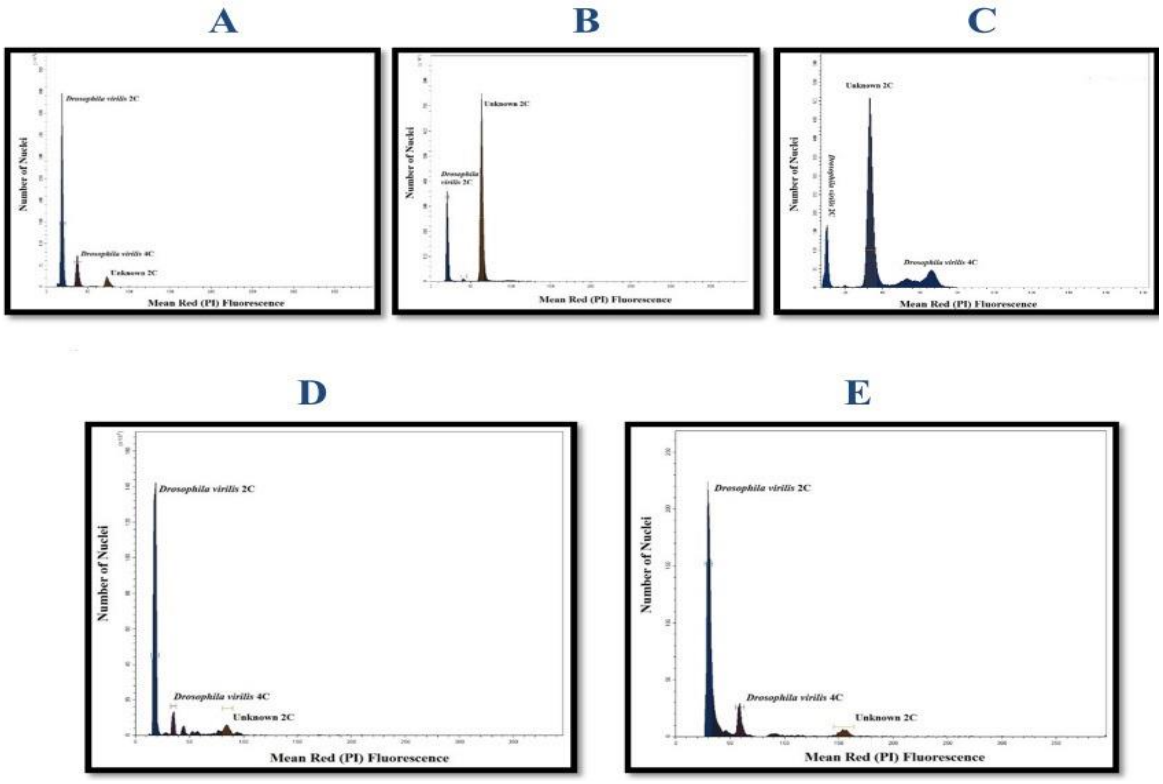

Fig (3) Flow cytometry histogram showing the relative fluorescence of co-prepared 2C nuclei from female Drosophila virilis nuclei and the unknown or known Sarcophagidae. A) at 606 Mbp B) at 1045 Mbp C) at 1105 Mbp D) at 1552.17 Mbp E) at 1734.88 Mbp. The genome size of Sarcophagidae is calculated as the ratio of the mean fluorescence of the 2C Sarcophagidae sp. to the mean fluorescence of the 2C *D. virilis* nuclei times the 1C amount of DNA in the latter (328 Mbp). A small peak of 4C nuclei at fluorescence channel 900 illustrates the important point that most nuclei from the head of both the sample and standard are 2C. There are extra peaks in many that can be identified as part of the challenge when the expected position of the 2C peak is unknown.

4. Discussion

Methods for genome size estimates via flow cytometry have traditionally called for live arthropod material [3]. This proved to be severely limiting for samples captured in the field as the majority of arthropods do not survive multiple days in vials.

Equally problematic is that field collections need to be identified before dissection, and many species are lost in the time between capture and dissection. Ethanol storage is widely used in insect curation and storage for genetic research, but is not useful for genome size estimation.

However, preliminary tests with ethanol produced no reliable peak representative of the stained nuclei. In contrast, isolation of nuclei from fresh frozen material proved remarkably successful. Samples frozen directly in the field with dry ice can be stored at − 20°C in the lab and run over a period of several weeks, with no change in the estimated genome size. Material frozen in − 20°C and − 80°C showed no difference in the estimated genome size over this storage period. Even the large insects such as the Blattaria: *Periplenata americana* produce excellent results after freezing [8].

In Insecta, relationships between ploidy level and genome size are available for *Bacillus atticus carius* Brunner (Phasmida: Phasmatidae) for which the ratio of the genome size between diploid and triploid lines is 1.5, which is consistent with a proportional increase in ploidy and genome size [15] and also for *Thrips tabaci* Lindeman (Thysanoptera: Thripidae) populations which the ratio of the genome size between diploid and polyploid [10].

Two genome size estimates for *W. magnifica* and *S. agrystoma* could be explained in several ways. First, sexes can be dimorphic due to differences in sex chromosome sizes. For instance, male white-backed brown planthopper (WBPH) *Sogatella furcifera* and male planthoppers (small brown planthopper (SBPH) *Laodelphax striatellus* have a significantly smaller genome size than females [9]. In blow flies, [16] observed sexual dimorphism in genomes size for some species. Second, it may be possible for polyploidy to produce signals if they appear to form a polyploidy series, 2N, 4N with each genome size reduced by 40% following the ploidy event as in *W. magnifica*. In Insecta, relationships between ploidy level and genome size are available for *Bacillus atticus carius* Brunner (Phasmida: Phasmatidae) for which the ratio of the genome size between diploid and triploid lines is 1.5, which is consistent with a proportional increase in ploidy and genome size [15] and also for *Thrips tabaci* Lindeman (Thysanoptera: Thripidae) populations which has a similar ratio of the genome size between diploids and polyploids [10]. Here, sexual dimorphism and ploidy differences fit PCR results that confirm they are all the same species. However, covariance between traits in genome sizes might be linked to ecological and physiological differences among species, as in the *Udranomia kikkawai* complex in Costa Rica revealed to possess cryptic species [11].

5. Conclusion

There are differences in genome size both within and between species of Egyptian Sarcophagidae when immature samples are analyzed. There is a need to repeat the experiments with adults to see if there is sexual dimorphism in genome size, polyploidy, or cryptic species. Despite the remaining challenges to implementation, this study lays the foundation for flow cytometry / genome size based species identification in the Sarcophagidae for forensic, medical, and agricultural purposes.

6. Acknowledgements

This project was supported by the **Egyptian Cultural Affairs and Mission Sector** (Cairo, Egypt) in the form a Scholarship as a part of the joint supervision program for Ahmed El-Hefnawy. Thanks to **Renee Holmes** (PhD candidate at Texas A&M University) for her role in teaching Ahmed step by step instructions on how to use the flow cytometer and for extracting the results.

References

[1] J.H.Byrd and J.L.CastnerForensic entomology: the utility of arthropods in legal investigations. CRC, Boca Raton, FL, 2011.

[2] R.F.Dennoand W. R. Cothran, Competitive interaction and ecological strategies of sarcophagid and calliphorid ßies inhabiting rabbit carrion. Ann. Entomol. Soc. Am. PP.69: 109-113, 1976.

[3] R.DeSalle, T.R. Gregory and J.S. Johnston, Preparation of samples for comparative studies of arthropod chromosomes: visualization, in situ hybridization, and genome size estimation. Methods. Enzymol. PP.395:460 – 488, 2005.

[4] D.W.Galbraith , K. R. Harkins, J. M. Maddox, N. M. Ayres, D. P. Sharma and E. Firoozabady, Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science. PP.220: 1049–1051, 1983.

[5] B.Greenberg, Flies as forensic indicators. J. Med. Entomol, Vol.28, pp. 565-577, 1991.

[6] B.Greenberg and J.C.Kunich, Entomology and the law. Flies as forensic indicators. Cambridge University Press, Cambridge, United Kingdom, 2002.

[7] T.R.Gregory, Animal genome size database. (http://www.genomesize.com), 2011.

[8] S.J.Hanrahanand J. S. Johnston, New genome size estimates of 134 species of arthropods. Chromosome Res. PP.19:809 – 823. DOI 10.1007/s10577-011-9231-6, 2011.

[9] He, K., K. Lin, G. Wang and F. Li, Genome Sizes of Nine Insect Species Determined by Flow Cytometry and k-mer Analysis. Front. Physiol. PP.7:569, 2016.
https://doi.org/10.3389/fphys.2016.00569.

[10] A.L.Jacobson , J. S. Johnston, D. Rotenberg, A. E. Whitfield, W. Booth, E. L. Vargo and G. G. Kennedy. Genome size and ploidy of Thysanoptera. Insect Molecular Biology. Vol.22(1), pp.12–17, 2013.

[11] D.H.Janzen, , J. M. Burns, Q. Cong, W. Hallwachs, T. Dapkey, R. Manjunath, M. Hajibabaei, P. D. N. Hebert, and N. V. Grishin, Nuclear genomes distinguish cryptic species suggested by their DNA barcodes and ecology. Proc Natl Acad Sci U S A. Vol.114(31): 8313–8318. doi: 10.1073/pnas.1621504114, 2017.

[12] D.H.Lunt, D. X. X. Zhang, J. M. Szymura, G. M. Hewitt, O. Hewltt, and D. X. Zhang. The Insect Cytochrome Oxidase I Gene: Evolutionary patterns and conserved primers for phylogenetic studies. Ins. Mol. Biol. Vol.5(3), pp. 153–165, 1996.

[13] D.Martín-Vega, M.J.R. Hall and T.J.Simonsen, Resolving confu-sion in the use of concepts and terminology in intrapuparial development studies of cyclorrhaphous Diptera. J. Medical Entomology, vol.53, pp.1249–1251, 2016.

[14] Martinez-Sanchez, A., K. E. Smith, S. Rojo, M. A. Marcos-Garcia, and R. Wall. Geographic origin affects larval competitive ability in European populations of the blow fly, *Lucilia sericata*. Entomol. Exp. Appl. PP.122: 93-98, 2007.

[15]B.B.Normark, Polyploidy of parthenogenetic *Aramigus tessellatus* (Say) (Coleoptera: Curculionidae). Coleopt. Bull. PP.50: 73–79, 1996.

[16] C.J.Picard , J.S.Johnston and A.M.Tarone , Genome sizes of forensically relevant Diptera. Journal of medical entomology; Vol.49(1), pp.192-7, 2012.

[17] C.J.Picard, and J. D. Wells, A test for carrion fly full siblings: a tool for detecting postmortem relocation of a corpse. J. Forensic Sci. Vol.57(2):535-538, 2012.

[18] F.A.Sperling, G. S. Anderson, and D. A. Hickey, A DNA-based approach to the identification of insect species used for postmortem interval estimation. J. Forensic Sci. Vol.39, pp.418-427, 1994.

[19] K.Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28: 2731-2739, 2011.

[20] A.M.Tarone, and M. R. Sanford,Is PMI the Hypothesis or the Null Hypothesis? J. of Med. Entomol. Vol.54 (5), pp.1109–1115, 2017.

[21] Wells, J. D., and F. A. Sperling. 2000. A DNA-based approach to the identification of insect species used for postmortem interval estimation and partial sequencing of the cytochrome oxidase b subunit gene I: a tool for the identification of European species of blow flies for postmortem interval estimation. J. Forensic Sci. 45: 1358-1359.

[22] Wells, J. D., and F. A. Sperling. 2001. DNA-based identification of forensically important Chrysomyinae (Diptera: Calliphoridae). Forensic Sci. Int. 120: 110-115.

[23] J.D.Wells, T. Pape, and F. A. Sperling. DNA-based identification and molecular systematics of forensically important Sarcophagidae (Diptera). J. Forensic Sci. Vol.46, pp. 1098-1102, 2001.