A Review on Phylogenetic Analysis Of Forensically Important Flies (DIPTERA)

Manish Sharma[1], Devinder Singh[2]
[1, 2] Department of Zoology and Environmental Sciences
[1, 2] Punjabi University, Patiala, India

I. HISTORICAL BACKGROUND

Phylogenetics is the study of evolutionary relationships. Phylogenetic analysis is the means of inferring or estimating these relationships. The evolutionary history inferred from phylogenetic analysis is usually depicted as branching, treelike diagrams that represent an estimated pedigree of the inherited relationships among molecules (gene trees), organisms, or both. Phylogenetics is sometimes called cladistics because the word clade, a set of descendants from a single ancestor, is derived from the Greek word for branch. However, cladistics is a particular method of hypothesizing about evolutionary relationships [1]. German entomologist Willi Henning published a book Grundzuge einer Theorie der Phylogenetischen systematic in 1950 [2] and this led to a wide spread exercise of this lucid method.

Haeckel [3] is usually regarded as the first evolutionary biologists to publish an explicit phylogenetic tree of real organism although the honor should probably go to Lamarck. The phylogenetic tree means a diagram explicitly denoted as the evolutionary relationships among diverse forms. However, diagrams which look all intents and purposes like phylogenetic trees have been around at least since Agassiz [4] and Peterson [5] pointed out that they look strikingly like those of later authors such as Romer [6].

Phylogenetic trees have been depicted in various forms for various purposes. Haeckel's [3] diagram was actually a tree. Romer's [6] diagram of vertebrates incorporates an absolute time frame and an estimate of the numbers of species of each group at any one time. Milne and Milne [7] tree of caddis fly is three dimensional and incorporates such factors as habitat and casing contractions. Phylogenetic trees may have characters placed on them or they may be bare. In short, phylogenetic trees come in many shapes and forms, each with a slightly different emphasis on the subject [1].

The first accounts of dipteran phylogeny appeared more than one and a half century ago [8]. A large number of studies followed, which derived phylogenetic evidence from cytology, paleontology, or morphology [9-11].

Species identification by phylogenetic analysis can be extremely accurate, provided the initial database is large enough to cover all unknown species, or the unknown species can be identified to a level where they will be covered by the analysis [12]. The accuracy of this identification technique makes it ideal for use in legal forensic diagnostics, where the identification may undergo close legal scrutiny. It has been used for this purpose, most notably to identify flies (Diptera) of forensic importance during legal proceedings. However, the time involved in making this robust identification is not well suited to large scale diagnostic identification, where high accuracy is of paramount importance [12-15].

Phylogenetic trees are diagrams depicting the evolutionary descent of whole organisms or groups of whole organisms. Reconstruction of phylogenetic trees is one of the most fundamental problems in computational biology. A phylogenetic tree reveals the evolutionary relationship among a given set of species. The basic hypothesis is that all organisms on Earth are evolutionarily related via a common ancestor. Notably, genes can be diverged by either gene duplication (paralog) or speciation (ortholog). For the construction of phylogenetic trees of species it must be based on orthologs. We also need to emphasize that even the phylogenetic tree that best explains the sequence data of a group of species does not necessarily represent the true phylogenetic tree of a species due to the processes of gene duplication, loss and lineage sorting [16]. In general, phylogenetic tree construction methods can be classified into four categories: distance based methods, maximum parsimony methods, maximum likelihood schemes and maximum compatibility methods. Distance based methods include UPGMA and neighbor joining algorithms [17].

II. PHYLOGENETIC ANALYSIS OF FLIES

Wallman and Donnellan [18] sequenced partial of mitochondrial COI and COII and investigated forensically important immature blowflies from the western Australia. Analysis of COI and COII sequences revealed abundant phylogenetically informative nucleotide divergence of sister species, the data could not distinguish among taxa form the same species group i.e., the species within the Chrysomya augur and Chrysomya stygia group.

Wells et al. [19] attempted a study to distinguish between Chrysomya chloropyga and Chrysomya putoria using mitochondrial DNA (mtDNA) sequence data from a 593-bp region of the gene for cytochrome oxidase subunit one (COI). Twelve specimens from each species yielded a total of five haplotypes, none being unique to C. putoria. Therefore it was not possible to distinguish between the two species using this locus. Maximum parsimony analysis indicated paraphyletic C. chloropyga mtDNA with C. putoria nested therein. Based on these and previously published data, we infer that C. putoria diverged very recently from C. chloropyga.

Zehner et al. [14] demonstrated phylogenetic analysis of 12 species of Sarcophagidae using COI and ND5 genes and found that these genes are suitable for the correct identification of flesh flies of German origin.

Nelson et al. [18] The utility of cytochrome oxidase I (COI) DNA barcodes for the identification of nine species of forensically important blowflies of the genus Chrysomya (Diptera: Calliphoridae), from Australia, was tested. A 658-bp fragment of the COI gene was sequenced from 56 specimens, representing all nine Chrysomya species and three calliphorid outgroups. Nucleotide sequence divergences were calculated using the Kimuratwo-parameter distance model and a neighbour-joining (NJ) analysis was performed to provide a graphic display of the patterns of divergence among the species. All species were resolved as reciprocally monophyletic on the NJ tree. Mean intraspecific and interspecific sequence divergences were 0.097% (range 0 – 0.612%, standard error [SE] = 0.119%) and 6.499% (range 0.458 – 9.254%, SE = 1.864%), respectively. In one case, a specimen that was identified morphologically was recovered with its sister species on the NJ tree.

Park et al. [19] reported the full-length sequences of the COI gene of four Calliphorinae fly species collected in Korea (five individuals of Calliphora vicina, five Calliphora lata, four Triceratopyga calliphoroides and three Aldrichina grahami). Each COI gene was amplified by polymerase chain reaction and directly sequenced and the resulting nucleotide sequences were aligned and analyzed by MEGA4 software. The results indicate that COI nucleotide sequences can be used to distinguish between these four species. Our phylogenetic result coincides with recent taxonomic views on the subfamily Calliphorinae in that the genera Aldrichina and Triceratopyga are nested within the genus Calliphora.

Tuorle et al. [20] sequenced complementary nuclear (28S rRNA) and mitochondrial (COI) genes from blowflies that phenotypically resembled Lucilia cuprina (W.), Lucilia sericata (Meigen) or exhibited characters of both species. The aim was to test a long held hypothesis that these species hybridize under natural conditions in South Africa [21]. Blowflies were obtained predominantly from the Cape Town metropolitan area, but reference samples were acquired for L. sericata from Pretoria. Several L . cuprina -like flies were shown to possess a conflicting combination of nuclear and mitochondrial genes that has also been seen in Hawaiian specimens.

Bajpai and Tewari [22] showed the phylogenetic relationships of 5 species of Sarcophagidae on the basis of 292 bp region of COI and ND5 genes of mtDNA from Indian origin. The minimum evolution trees were constructed with sequences of both the genes for all the species and found to be useful for the identification of these species in India.

DeBry et al. [23] provided the first DNA data for Lucilia cuprina from North America, including portions of both the mitochondrial COI gene and the nuclear 28S rRNA gene. With the new data, L. cuprina remains monophyletic for 28S but paraphyletic with respect to Lucilia sericata for COI. However, they find that all flies that are identified as L. cuprina by morphology and have L. sericata-like mtDNA form a distinctly monophyletic mtDNA clade. This clade may possibly have originated by hybridization between L. cuprina and L. sericata, but its wide geographic distribution strongly suggests a singular origin as opposed to repeated incidents of hybridization. The phylogenetic results strongly support the hypothesis that L. cuprina and L. sericata can be discriminated using mtDNA sequence data. They find that a fragment of COI spanning approximately 1200 base pairs is sufficient to discriminate between the two species with greater than 95% bootstrap support.

Similarly, Guo et al. [24] used 272 bp region of COI gene of four species of Sarcophagidae for phylogenetic analysis. Parsimonious tree showed high bootstrap values (100%) which provide an indication of good percentage support for the grouping nodes of S. peregrina, S. similis and S. albiceps. Within the S. dux species, 2 specimens from Yongzhou and 1 specimen from Changsha were clustered together with the support value of 100%, as they were all obtained from adjoining areas.

Guo et al. [25] analyzed 55 COI gene sequences from 7 species of Chinese Sarcophagidae through NJ tree by using Kamura-2-Parameter algorithm. They showed that at the species level, the high bootstrap values (100%) provide robust support for the monophyly of P. gravelyi and P. similis. Within P. dux, two specimens clustered together with a supporting bootstrap of 99%, and other specimens were sister to them with a supporting value of 85%. Within P. albiceps,

seven specimens clustered together with a weak supporting bootstrap of 69%, and the specimen from Changsha (GenBank accession number JF416549) was sister to them with a supporting value of 100%. However, different branches were formed sharing low supporting values for B. peregrina, S. javanica and B. africa. At the genus level, P. albiceps and P. dux cluster together, whereas B. peregrina, S. javanica and P. similis cluster together indicating the ability of this shorter COI fragment to identify the species from the same genus was not as efficient as that of the longer fragments.

Subsequently, Guo et al. [26] again used 272 bp region of COI and 16S rRNA gene of 4 species of Sarcophagidae for phylogenetic analysis. The phylogeny of COI gene of these flies was clearly separated into four genetic clades (A–D). Clade A, which comprised of 9 S. dux specimens, was supported by high bootstrap value. Clade B was known as the albiceps group as it only contained species S. albiceps. Within clade C, the peregrina group, two different branches were formed sharing a supporting value of 57%. Four S. melanura specimens were united under clade D as the melanura group. Clade A and clade B belonged to the genus Parasarcophaga, clustered together with a supporting bootstrap of 60%. Guo et al. [27] again studied five species of Sarcophagidae by constructing UPGMA trees and showed that all the species correctly assigned into five groups with monophyletic separation in both COII as well as 16s rRNA UPGMA trees.

Meiklejohn et al. [28] analyzed phylogenetically 16 Australian species of genus Sarcophaga. The intraspecific variation within the non-monophyletic species of S. dux as depicted from the NJ tree was 4.658%. Based on 'barcode' thresholds, this value is indicative of two distinct species which is portrayed graphically by separate clusters in the NJ tree. The range of interspecific variation for directly comparable Australian Sarcophagidae, as determined from examination of the NJ tree, was 6.658–8.983% (SE=0.653%), except the 2.891% interspecific variation noted between S. megafilosia and S. meiofilosia suggesting that these morphologically distinct specimens could have diverged very recently or possibly belong to the same species.

Jordaens et al. [29] sequenced the mitochondrial COI gene of 126 specimens of 56 European Sarcophaga species. They showed that the identification success using a mini barcode region of 127 bp was very low (80.7–82.5 %) and the use of this region is not recommended as a species identifier. In contrast, identification success was very high using the standard barcode region (658 bp) or using the entire COI region (1,535 bp) (98.2–99.3%).

Kavitha et al. [30] evaluated the usefulness of the molecular and morphological approach in species identification using morphologically identified 10 blow fly larvae sampled from 10 different crime scenes in Malaysia. Their study involves the sequencing of a total length of 2.2 kilo base pairs of the mitochondrial COI, COII and t-RNA leucine genes. They concluded that morphology and molecular identification were in concordance in eight cases (80%). In the case 7, the recovered larvae were identified morphologically as C. pungius but later molecularly confirmed to be C. megacephala in NJ tree during the phylogenetic analysis. They further revealed that in case 1, the 3rd instar maggot which was identified as H. ligurriens morphologically but turned out to be an unidentified species of blow fly upon phylogenetic analysis of the entire cytochrome oxidase gene sequences. The BLAST results for case 1 showed a 94% similarity to L. cuprina.

Stamper et al. [31] inferred phylogenetic relationships using COI, COII and ND4 genes of mtDNA among species and genera of the subfamily Sarcophaginae by Bayesian Inference and Maximum Likelihood methods. Their results were consistent with monophyly of the subfamily Sarcophaginae (bootstrap support 93%), as well as with monophyly of several genera within the Sarcophaginae (including Sarcophaga with bootstrap support of 97%). They also found support for a sister group relationship between Ravinia Robineau-Desvoidy and Oxysarcodexia Townsend, which has been hypothesized by past authors on the basis of morphological similarities, although this was supported only in the Bayesian analyses (posterior probability 0.81–0.98).

Roziah et al. [32] used a molecular approach to examine the use of 12S & 16S ribosomal mitochondrial DNA (1172 bp) and Internal Transcribed Spacer (ITS) region of nuclear DNA (~1500 bp) to discriminate four species of the subgenus Boettcherisca namely, Boettcherisca javanica, B. peregrina, B. karnyi and B. highlandica. Neighbour-Joining phylogenetic tree was generated for each gene. Interspecific values calculated using the Kimura-2-parameter distance model were in the low range of 0.5%-1.3% and 0.5%-1.9% for 12S & 16SrDNA and ITS region respectively. However, identifications of B. karnyi and B. peregrina using ITS; and B. peregrina using 12S & 16S alone can be ambiguous. Therefore phylogenetic tree analyses of both genes showed a likely for these specimens to be distinguished and confirmed the potential of these genes as specific sarcophagid identification markers.

Yusseff-Vanegas and Agrarsson [33] presented the first complete phylogeny of Cochliomyia including numerous specimens per species, collected from 13 localities in the

Caribbean. Four genes, the mitochondrial COI and the nuclear EF-1α, 28S rRNA, and ITS2, were analyzed. While we found some differences among gene trees, a concatenated gene matrix recovered a robustly supported monophyletic Cochliomyia with Compsomyiops Townsend as its sister group and recovered the monophyly of Cochliomyia hominivorax, Cochliomyia macellaria and Cochliomyia minima. Their results support a close relationship between Cochliomyia minima and Cochliomyia aldrichi. However, They found Cochliomyia aldrichi containing Cochliomyia minima, indicating recent speciation, or issues with the taxonomy of the group and provide basic information on habitat preference, distribution and feeding habits of Cochliomyia minima and Cochliomyia aldrichi that will be useful for future forensic studies in the Caribbean.

Williams et al. [34] partially sequenced the 28S rRNA, COI and Period genes of 14 species of Lucilia and Hemipyrellia and analysed together with sequences of 11 further species from public databases. The molecular data confirmed molecular paraphyly in three species-pairs in Lucilia that hamper barcode identifications of those six species. Lucilia sericata and Lucilia cuprina were confirmed as mutual sister species. The placements of Dyscritomyia and Hypopygiopsis were ambiguous, since both made Lucilia paraphyletic in some analyses. Recognising Hemipyrellia as a genus consistently left Lucilia s.l. paraphyletic, and the occasionally-recognised (sub)genus Phaenicia was consistently paraphyletic, so these taxa should be synonymised with Lucilia to maintain monophyly.

REFERENCES

[1] Wiley, E.O. “Phylogenetics, the theory and practice of phylogenetic systematic”. John Wiley and Sons, New York, 1981.

[2] Henning, W. “Grundzuge einer Theorie der phylogenetischen Systematik”. Deutscher Zentraverlag, Berlin, 1950, pp, 370.

[3] Haeckel, E. “Generelle Morpholigie der Organismen”, II. Syst. Zool., 25: 1866, pp.168-173.

[4] Agassiz, L. (1844) Recherches see les Poissons Fossiles. Neuchatel, 5 vols., With supplement.

[5] Peterson, C. “The contribution of paleontology to teleostean phylogeny. In: Major patterns in vertebrate evolution. Hecht, Goody and Hecht (Eds.), Plenum Press, New York, 1977, pp. 579-643.

[6] Romer, A.S. “Vertebrate paleontology”, University of Chicago Press, Chicago, Illinois, 1966, pp. 468.

[7] Milne, M.J. and Milne, L.J. “Evolutionary trends in caddis worm case construction”. Ann. Entomologic. Soc. Am., 32: 1939, pp.533- 542.

[8] Lameere, A. “Notes pour la classification des Dipteres”. Memoires de la Societe Entomologique de Belgique, 12: 1906, pp.105-140.