**Article**

### Comparative Mitogenomics of Flesh Flies: Implications for Phylogeny

Jin Shang 1,†, Wentian Xu 1,†, Xiaofang Huang 1, Dong Zhang 1,* ©, Liping Yan 1,* and Thomas Pape 2

1 School of Ecology and Nature Conservation, Beijing Forestry University, Beijing 100083, China
2 Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark
* Correspondence: zhangdong_bjfu@bjfu.edu.cn (D.Z.); yanlp@bjfu.edu.cn (L.Y.)
† These authors contributed equally to this work.

**Simple Summary:** Flesh flies (Diptera: Sarcophagidae) make up the second-largest family of Oestroidea. They are well known for their veterinary, forensic, and medical importance due to their extremely diverse feeding habits, making them a hotspot in dipterology research. Therefore, efforts have been devoted to accumulating the mitochondrial genomes of Sarcophagidae. However, the mitogenome of flesh flies has mostly been sequenced for the genus *Sarcophaga* of the subfamily Sarcophaginae, and the other two subfamilies, Miltogrammini and Paramacronychiinae, have barely been touched. Here, we sequenced the mitochondrial genomes of five sarcophagid species from four genera representing all three subfamilies and investigated the phylogeny and evolution of flesh flies from the perspective of the mitogenome via comprehensive comparative analyses employing all mitogenomic data. This study will broaden our knowledge of flesh flies and make contributions to forensics, veterinary science, entomology, and ecology.

### Abstract

Flesh flies (Diptera: Sarcophagidae) represent a rapid radiation belonging to the Calyptratae. With more than 3000 known species, they are extraordinarily diverse in terms of their breeding habits and are therefore of particular importance in human and veterinary medicine, forensics, and ecology. To better comprehend the phylogenetic relationships and evolutionary characteristics of the Sarcophagidae, we sequenced the complete mitochondrial genomes of five species of flesh flies and performed mitogenomic comparisons amongst the three subfamilies. The mitochondrial genomes match the hypothetical condition of the insect ancestor in terms of gene content and gene arrangement. The evolutionary rates of the subfamilies of Sarcophagidae differ significantly, with Miltogrammini exhibiting a higher rate than the other two subfamilies. The monophyly of the Sarcophagidae and each subfamily is strongly supported by phylogenetic analysis, with the subfamily-level relationship inferred as (Sarcophaginae, (Miltogrammini, Paramacronychiinae)). This study suggests that phylogenetic analysis based on mitochondrial genomes may not be appropriate for rapidly evolving groups such as Miltogrammini and that the third-codon positions could play a considerable role in reconstructing the phylogeny of Sarcophagidae. The protein-coding genes ND2 and ND6 have the potential to be employed as DNA markers for species identification and delimitation in flesh flies.

**Keywords:** mitogenome; sarcophagidae; phylogeny; evolution

### 1. Introduction

Sarcophagidae, or flesh flies, which comprise more than 3000 known species, represent one of the most species-rich families of Calyptratae [1], and they have recently been the focus of several comprehensive studies [2–7]. They are widely distributed on all continents except Antarctica [2,4] and are divided into three subfamilies, namely, Miltogrammini, Paramacronychiinae, and Sarcophaginae [4]. Flesh flies are of great ecological and economic importance because of their diverse feeding habits. Necrophagous species breeding in...
animal corpses and excreta play an essential role in forensic cases [8–10], and agents of myiasis have considerable importance for human and veterinary medicine [11,12]. Lower Miltogramminae are sarcosaprophagous while the higher Miltogramminae are kleptoparasites of Hymenoptera, with larvae developing in the nests of their hosts, feeding on the food stored for the wasp or bee progeny [6,13,14]. Paramacronychiinae are mostly parasitoids or predators of invertebrates, but the genera Sarcophila and Wohlfahrtia contain scavengers and agents of myiasis [15]. For example, Agria contains parasitoids and predators of insects [16,17], and some species of Wohlfahrtia produce traumatic myiasis in humans and livestock [18,19]. Sarcophagines are the largest subfamily within Sarcophagidae. They are biologically diverse flies and are parasitoids or predators of invertebrates and scavengers of carrion and dung or feces [2,4].

The mitochondrial genome of insects is a small-sized organelle genome that is inherited maternally and seldom recombined [20]. Due to its small size, large number of copies, conserved genome components, fast evolutionary rate, and uncommon recombination, the mitochondrial genome has become a formidable tool in phylogenetic and evolutionary studies of various insect groups [20–28]. In comparative studies, mitogenome features such as compositional bias and substitutional rate variation provide essential information for the reconstruction of phylogeny as a basis for tracking evolution among insects [24,26,29–32]. Despite the fast accumulation of sequenced mitogenomes generated by high-throughput sequencing [33], our understanding of mitogenomic evolution in insects remains fragmentary.

To improve our understanding of the diversity in the mitochondrial genome and the evolution of mitochondrial genetics of Sarcophagidae, we sequenced the mitochondrial genomes of five sarcophagid species belonging to four genera and representing all three subfamilies. The genomic structure, base composition, AT content, substitutional and evolutionary rates, and the compositional heterogeneity of the mitogenomes were investigated, and the phylogeny of Sarcophagidae was reconstructed based on all mitogenomes available for flesh flies from GenBank.

2. Materials and Methods

2.1. Sampling and Taxa Identification

Specimens of Taxigramma karakulensis (Miltogramminae), Wohlfahrtia balassogloi and Wohlfahrtia fedtschenkoi (Paramacronychiinae), and Blaesoxipha lapidosa (Sarcophaginae) from the Kalamaili Nature Reserve in Xinjiang and A. mihalyii (Paramacronychiinae) from Songshan in Beijing were collected with a hand net and identified following Ye [34] and Xue et al. [35] (Table S1). The specimens were stored at −20 °C in 96% ethanol until DNA was extracted. Classification followed the most recent world catalogue [4] with subsequent updates [2,3,7].

2.2. DNA Extraction, Sequencing and Assembly

A Tiangen DNeasy Blood and Tissue Kit (Tiangen, Beijing, China) was used to extract genomic DNA from muscle tissue of each specimen. Among them, Blaesoxipha lapidosa and Taxigramma karakulensis were sequenced in combination, while Agria mihalyii, Wohlfahrtia balassogloi, and Wohlfahrtia fedtschenkoi were sequenced separately. A segment of mitochondrial gene COI was amplified and sequenced for each species by polymerase chain reaction (PCR) with universal primers as the “reference sequence” as described in Yan et al. [36]. Subsequently, the genomic DNA Library was prepared pooling the extracted genomic DNA and sequenced with a NovaSeq 6000 (PE150; Illumina, San Diego, CA, USA) platform. Trimmomatic [37] was used to trim raw data as described in Yan et al. [2] before assembly using IDBA_UD v 1.1.1 [38]. The assembly proceeded with similarity set to be 98%. The mitochondrial genomes were then extracted by the BLAST program [39] with the COI sequences as the “bait” [23].
2.3. Mitogenome Annotation and Sequence Analyses

Annotation of protein-coding genes (PCGs) and ribosomal RNA (rRNA) genes was performed using Geneious by comparing to homologous genes in other Calyptratae flies. The MITOS online (http://mitos2.bioinf.uni-leipzig.de/index.py, accessed on 5 May 2022) server and the tRNAScan-SE (http://lowelab.ucsc.edu/tRNAscan-SE/, accessed on 5 May 2022) online search server were used to determine the transfer RNA (tRNA) genes. A circular map of the mitochondrial genome was generated using the online server CGView (https://proksee.ca/, accessed on 7 May 2022).

All sarcophagid mitochondrial genomic data from GenBank were collected and used for comparative mitogenomic analyses together with the five newly sequenced mitogenomes in the present study (Table 1). We used MEGA X [40] to examine the base composition and codon usage of PCGs. To assess mitochondrial strand asymmetry of 13 protein-coding genes in all sarcophagid mitochondrial genomes, the AT skew and the GC skew were calculated with the following formulae: AT skew = (A − T)/(A + T) and GC skew = (G − C)/(G + C).

Table 1. Taxonomic information and GenBank accession number of mitochondrial genomes used in the study.

| Family   | Subfamily          | Genus            | Subgenus  | Species        | Accession No. |
|----------|--------------------|------------------|-----------|----------------|---------------|
| Sarcophagidae | Miltogramminae   | Mesomelena       | mesomelena| KY003227       |
|          | Miltogramma       | oestracea        |           | MW556020       |
|          | Phrosinella       | nasuta           |           | MW546340       |
|          | Protomiltogramma  | cincta           |           | NC_063942      |
|          | Sarcotachina      | subcylindrica    |           | MW556014       |
|          | Senotamina        | albifrons        |           | MW556016       |
|          | Sphecapatoclea    | sp.              |           | MW556015       |
|          | Taxigramma        | karakulens*      |           | ON375459       |
|          | Paramacronychiinae| Agria            | mihalyii* | ON360967       |
|          |                    | honshuensis      |           | MW556017       |
|          |                    | latifrons        |           | MW556018       |
|          |                    | balassoglai*     |           | ON411642       |
|          |                    | cheri            |           | MW556019       |
|          |                    | fedtschenkoi*    |           | ON411643       |
|          |                    | magnifica        |           | KU578263       |
| Sarcophaginae | Miltogramminae   | Mesomelena       | mesomelena| KY003227       |
|          |                    | oestracea        |           | MW556020       |
|          |                    | nasuta           |           | MW546340       |
|          |                    | cincta           |           | NC_063942      |
|          |                    | subcylindrica    |           | MW556014       |
|          |                    | albifrons        |           | MW556016       |
|          |                    | sp.              |           | MW556015       |
|          |                    | karakulens*      |           | ON375459       |
|          |                    | mihalyii*        |           | ON360967       |
|          |                    | honshuensis      |           | MW556017       |
|          |                    | latifrons        |           | MW556018       |
|          |                    | balassoglai*     |           | ON411642       |
|          |                    | cheri            |           | MW556019       |
|          |                    | fedtschenkoi*    |           | ON411643       |
|          |                    | magnifica        |           | KU578263       |
|          |                    | lapidosa*        |           | OM640654       |
|          |                    | latisterna       |           | KT272844       |
|          |                    | boettcherisa     |           | KT272848       |
|          |                    | diminuta         |           | MT017719       |
|          |                    | graciliformes    |           | MT017712       |
|          |                    | josephi          |           | MT017711       |
|          |                    | ptergyota        |           | MK820722       |
|          |                    | africa           |           | KM881633       |
|          |                    | formosensis      |           | MF688648       |
|          |                    | peregrina        |           | KF921296       |
|          |                    | melanura         |           | KP091687       |
|          |                    | depressifrons    |           | MT017709       |
|          |                    | pauciseta        |           | NC_053729      |
|          |                    | plotnikovi       |           | MT017720       |
|          |                    | shnitnikovi      |           | MT017730       |
|          |                    | tsinanensis      |           | NC_057991      |
|          |                    | cetu             |           | MT017725       |
|          |                    | schuetzei        |           | MT017726       |
|          |                    | crassipalpis     |           | KC005711       |
|          |                    | raficornis       |           | MF937749       |
Table 1. Cont.

| Family          | Subfamily | Genus   | Subgenus | Species          | Accession No. |
|-----------------|-----------|---------|----------|------------------|---------------|
| Liosarcophaga   |           | aegyptica |          | MT017732         |
| Liosarcophaga   |           | brevicornis |        | NC_047404       |
| Liosarcophaga   |           | dux      |          | MH879759        |
| Liosarcophaga   |           | jacobsoni |          | MT017723        |
| Liosarcophaga   |           | kanoi    |          | MT476487        |
| Liosarcophaga   |           | porstchinskyi |      | KM288750        |
| Liosarcophaga   |           | tuberosa |          | MK820723        |
| Neobellieria    |           | bullata  |          | KT272859        |
| Pandelleisca    |           | kawaynensis |      | MT017713        |
| Pandelleisca    |           | pingi    |          | MT017728        |
| Pandelleisca    |           | scapariiformis |    | MT476486        |
| Pandelleisca    |           | similis  |          | KM257431        |
| Parasarcophaga  |           | albiceps |          | KT444443        |
| Parasarcophaga  |           | macroauriculata |  | MT017718        |
| Parasarcophaga  |           | misera   |          | MF133500        |
| Phallocheta     |           | minor    |          | MT017727        |
| Pseudothyrsocnema |         | caudagalli |        | MK820721        |
| Robineauella    |           | anchoriformis |    | MT017716        |
| Sarcophaga      |           | carnaria |          | MT017710        |
| Sarcorohdendorfia |        | antilope |          | MH540748        |
| Sarcorohdendorfia |        | gracilis |          | MW531675        |
| Sarcorohdendorfia |        | impatiens |        | JN859549        |
| Seniorwhitea    |           | princeps |          | MH981944        |
| Sinonipponia    |           | hervebazini |     | MT017708       |
| Thyrsocnema     |           | kentjena |          | MT017714        |
| Ravinia         |           | pernix   |          | KM676414        |
| Oxysarcodexia   |           | avuncula |          | MH879754        |
| Oxysarcodexia   |           | terminalis |       | MH879757       |
| Oxysarcodexia   |           | thornax  |          | MH879765        |
| Oxysarcodexia   |           | varia    |          | MH879764        |
| Peckia          |           | australis |         | MH879762        |
| Peckia          |           | collusor |          | MH879763        |
| Peckia          |           | resona   |          | MH879761        |
| Calliphoridae   |           | vomitoria |        | KT444440        |
| Anthomyiidae    |           | antiqua  |          | NC_028226       |
| Fanniidae       |           | scalaris |          | MT017706        |
| Muscidae        |           | domestica |        | KM200723        |

* The sequences for species with names in bold were generated in this study.

The nonsynonymous substitutions (Ka) and synonymous substitutions (Ks) of the PCGs were calculated using DnaSP v6.12.03 (Barcelona, Spain) [41]. The genetic distances were calculated using MEGA X with a Kimura 2-parameter model. The Xia model test in DAMBE software [42] was used to evaluate the substitution saturation of each codon position. Bowker’s test of symmetry was performed with SymTest v2.0.47 (San Francisco, CA, USA) [43] to investigate the heterogeneous sequence divergence within the dataset, and heatmaps were generated from the calculated p-values.

2.4. Phylogenetic Analysis

The phylogeny reconstruction was conducted using all available sarcophagid mitogenomes (Table 1). Fannia scalaris (Fanniidae), Musca domestica (Muscidae), Calliphora vomitoria (Calliphoridae), and Delia antiqua (Anthomyiidae) were included as outgroups, and F. scalaris was selected to root the tree.

Multiple sequence alignments (MSAs) of all of the protein-coding and rRNA genes were obtained by aligning each gene using the online version of MAFFT v7.453 (Osaka, Japan) [44]. Four matrices were then concatenated using SequenceMatrix v1.8.1 (Science Drive, Singapore) [45]: (i) the matrix PCGsRNA containing all 13 PCGs and 2 rRNA genes,
(ii) the matrix PCGs12 containing only the 1st- and 2nd-codon positions of the 13 PCGs, (iii) the matrix PCGs containing the 13 PCGs, and (iv) the matrix PCGs12rRNA containing the 1st- and 2nd-codon positions of the 13 PCGs and 2 rRNA genes. Maximum Likelihood (ML) analyses were inferred using IQ-TREE v2.1.2 [46] on CIPRES (Cyberinfrastructure for Phylogenetic Research) Science Gateway (available at https://www.phylo.org/, accessed on 13 June 2022), with each dataset partitioned by gene. The best tree was searched using the best models evaluated by the self-implemented ModelFinder [47]. Non-parametric replicate bootstrapping with 100 repetitions was subsequently performed to evaluate the branch support. The resulting trees were visualized using the online tool iTOL [48] (available at https://itol.embl.de/upload.cgi (accessed on 15 June 2022).

3. Results and Discussion

3.1. General Features of Sarcophagid Mitogenomes

The mitochondrial genomes of five species of Sarcophagidae were successfully sequenced for the first time. The entire mitogenomes of *A. mihalyii*, *B. lapidosa*, *T. karakulensis*, *W. balassogloi*, and *W. fedtschenkoi* were 16,375 bp, 15,454 bp, 16,543 bp, 16,589 bp, and 16,643 bp in length, respectively (Figure 1). The variation in length is mainly due to the different size of the control region, as previously observed between other Sarcophagidae species [36]. Each mitogenome encoded 37 typical insect mitochondrial genes, 13 protein-coding genes (PCGs), 2 ribosomal RNA (rRNA) genes, and 22 transfer RNA (tRNA) genes. A total of 23 genes, including 9 PCGs and 14 tRNA genes, were transcribed from the majority strand (J-strand), while 14 genes were transcribed from the minority strand (N-strand). The nine PCGs transcribed from the J-strand were ATP6, ATP8, CYTB, COI, COII, COIII, ND2, ND3, and ND6, and the other four (ND1, ND4, ND4L, and ND5) were transcribed from the N-strand. Both rRNA genes (lrRNA and srRNA) were transcribed from the N-strand (Figure 1). All genes were in the same gene order and orientation as the hypothetical ancestral mitogenome, consistent with other calyptrate flies [21,22,24,49,50].

The whole mitogenome for each of *A. mihalyii*, *B. lapidosa*, *T. karakulensis*, *W. balassogloi*, and *W. fedtschenkoi* was biased towards A and T, with the AT content being 74.43%, 75.17%, 79.15%, 76.02% and 75.93%, respectively. Other species of Sarcophagidae also have this typical composition (Figure 2a), and all of them represent a moderate level for calyptrate flies (Figure S1). By comparing the AT content of the different PCGs for all Sarcophagidae, ND6 (82.10%) had the highest average AT content, followed by ATP8 (81.64%) and ND4L (81.33%). The average AT content of COI (67.45%) and COIII (68.08%) were the lowest.

The nucleotide skew analysis revealed negative AT skews (Figure 2b), indicating that the overall strand was asymmetric with protein-coding genes of all sarcophagid species containing less adenine (A) than thymine (T). However, GC skews were close to zero (Figure 2c), suggesting that the content of guanine (G) and cytosine (C) was not significantly different. The correlations were calculated for the AT content versus the AT skew \(y = -0.0006x - 0.1109, R^2 = 0.0125\) and for the GC content versus the GC skew \(y = -0.0112x + 0.2844, R^2 = 0.6557\). The results showed that, with increasing AT content, the underabundance of A relative to T increased slightly. However, with increasing GC content, the overabundance of G relative to C gradually diminished and shifted to an overabundance of C relative to G when the GC content reached a level of about 25%.
Figure 1. *Agria mihalyii* (a), *Blaesoxipha lapidosa* (b), *Taxigramma karakulensis* (c), *Wohlfahrtia balassogloi* (d), and *Wohlfahrtia fedtschenkoi* (e). The strands are marked with arrows that indicate the direction of gene transcription. Gene lengths correspond to the length of their nucleotides. The IUPAC-IUB single-letter amino acid codes assign one-letter symbols to tRNA genes. Gene names were represented by their abbreviations: COI–COIII, cytochrome oxidase subunits 1–3; CYTB, cytochrome b; ND1-6 and ND4L, NADH dehydrogenase subunits 1–6 and 4L; ATP6 and ATP8, ATP synthetase subunits 6 and 8; lrRNA and srRNA, large and small ribosomal RNA subunits; CR, control region.
Insects 2022, 13, x FOR PEER REVIEW 7 of 16

75.17%, 79.15%, 76.02% and 75.93%, respectively. Other species of Sarcophagidae also have this typical composition (Figure 2a), and all of them represent a moderate level for calyptrate flies (Figure S1). By comparing the AT content of the different PCGs for all Sarcophagidae, ND6 (82.10%) had the highest average AT content, followed by ATP8 (81.64%) and ND4L (81.33%). The average AT content of COI (67.45%) and COIII (68.08%) were the lowest.

Figure 2. Nucleotide composition analysis of mitochondrial genomes from three subfamilies of Sarcophagidae: A + T percentage of the 13 protein-coding genes (a) and the corrections between A + T% vs. AT skew (b) and G + C% vs. GC skew (c) in the 13 protein-coding genes.

3.2. Protein-Coding Genes and Codon Usage

The PCGs of A. mihalyii, B. lapidosa, T. karakulensis, W. balassogloi, and W. fedtschenkoi were 11,184 bp, 11,179 bp, 11,157 bp, 11,179 bp, and 11,179 bp in length, respectively. For the five newly sequenced mitogenomes, all the PCGs started with the codon ATN (ATA/T/G). The complete termination codon TAA or TAG was utilized by most of the PCGs in these five mitogenomes, while the incomplete stop codon T was employed by COI, COII, ND4, and ND5. All three stop codons are common in the mitochondrial genomes of insects. Moreover, Ojala et al. [51] speculated that the incomplete stop codon T could be converted into the complete stop codon TAA by post-transcriptional polyadenylation.

The relative synonymous codon usage (RSCU) of each mitogenome was estimated (Figure S2). The third-codon position was more likely to be A or T than G or C. The most
frequently encoded amino acids in Sarcophagidae were Ala, Arg, Gly, Leu, Pro, Ser, Thr, and Val, which have the greatest RSCU values, and the most commonly used codons were UUA, GCU, GGA, and GUA.

3.3. Ribosomal, Transfer RNA Genes, and Control Region

All 22 tRNA genes were detected and found to be discontinuously scattered throughout the mitochondrial genome in the same location as hypothesized for the ancestral insect mitogenome. The size of each tRNA gene for these newly sequenced mitogenomes was hardly different, ranging from 62 bp to 72 bp (Table S2). The nucleotide composition of tRNA genes is biased toward A and T, with the AT content of each gene ranging from 76.04% (B. lapidosa) to 77.60% (T. karakulensis). The AT content of the tRNA genes was slightly higher than that of the protein-coding genes, and the combined tRNA genes exhibited positive AT skew and negative GC skew, except B. lapidosa (Table S3).

The two rRNA genes were located on the N-strand, with IrRNA between the tRNA1L and the tRNA2V genes, ranging from 1322 bp (W. fedtschenkoi) to 1329 bp (A. mihalyii and T. karakulensis), and srRNA located between the tRNA2V gene and the control region ranging from 783 bp (A. mihalyii) to 790 bp (W. fedtschenkoi) (Table S2). The AT content of the IrRNA genes ranged from 80.56% (W. fedtschenkoi) to 82.77% (T. karakulensis), while that of the srRNA genes ranged from 75.89% (W. balassogloi) to 79.97% (T. karakulensis). The AT skew and GC skew of IrRNA genes for the five species were negative and positive, respectively. The srRNA genes exhibited negative AT skew for all of the newly sequenced species except T. karakulensis (0.03), and positive GC skew was observed for all five species (Table S3).

The length of the control region of sarcophagid species varied considerably (A. mihalyii, B. lapidosa, T. karakulensis, W. balassogloi, and W. fedtschenkoi were 1545 bp, 650 bp, 1753 bp, 1789 bp, and 1828 bp in length, respectively), whereas there were few differences in the remaining region. Extrapolated from the information of these five species (Table S2), the length of the mitochondrial genome of sarcophagid species mainly depends on the size of the control region, which is consistent with other insects. The AT content in the control region ranged from 83.37% (A. mihalyii) to 89.85% (T. karakulensis), which is significantly higher than for the other regions of the mitogenomes, indicating a strong AT bias in this region (Table S3).

3.4. Evolutionary Rates and Heterogeneous Sequence Divergence

All three subfamilies have similar synonymous substitution rates (Ks), while the nonsynonymous substitution rates (Ka) of Miltogramminae were in all cases higher than those of Paramacronychiinae and Sarcophaginae (Figure 3). The Ka/Ks ratio (ω) is used for investigating the evolutionary rate. All 13 PCGs were suggested to be under strong purifying selection, with the ω of each gene calculated to be less than 1.0. The Ka/Ks ratio in Miltogramminae was higher than in the other two subfamilies, indicating an accelerated evolutionary rate in Miltogramminae.

Furthermore, the pairwise nonsynonymous (Ka) to synonymous (Ks) substitution ratio (ω) of all the PCGs in Sarcophagidae was analyzed, and the Ka/Ks ratio of ATP8 (ω = 0.286) and ND6 (ω = 0.203) was higher, indicating that these genes were under relatively little selection pressure (Figure 4). However, the Ka/Ks ratio of COI was the lowest (ω = 0.051), indicating a higher selection pressure for COI than for other genes. Based on the pairwise genetic distance calculation, ND6 (0.205), ND2 (0.175), and ND3 (0.174) had the highest values, while ND1 (0.105), ND4L (0.108), and ND4 (0.126) had the lowest.
Generally, the mitochondrial COI gene is the most extensively utilized molecular marker on the molecular level using appropriate molecular markers and DNA barcoding techniques. The taxonomic identification of flies is still mostly limited to morphology-based identification of adults, but the identification of species with complex morphological features or extreme similarity becomes quite tricky. However, this problem can be solved at the molecular level using appropriate molecular markers and DNA barcoding techniques. Generally, the mitochondrial COI gene is the most extensively utilized molecular marker for molecular identification and phylogenetic analysis among species [52–55]. In this study, COI was the most conserved, most slowly evolving, and least variable gene among all PCGs in the Sarcophagidae. If the COI gene is proven to provide low-resolution power, then alternative genes with sufficient size, relatively high Ka/Ks ratio, and rapid evolution could be used as potential DNA markers [52,56], notably ND2 and ND6, for species identification and delimitation in flesh flies.

The taxonomic identification of flies is still mostly limited to morphology-based identification of adults, but the identification of species with complex morphological features or extreme similarity becomes quite tricky. However, this problem can be solved at the molecular level using appropriate molecular markers and DNA barcoding techniques. Generally, the mitochondrial COI gene is the most extensively utilized molecular marker for molecular identification and phylogenetic analysis among species [52–55]. In this study, COI was the most conserved, most slowly evolving, and least variable gene among all PCGs in the Sarcophagidae. If the COI gene is proven to provide low-resolution power, then alternative genes with sufficient size, relatively high Ka/Ks ratio, and rapid evolution could be used as potential DNA markers [52,56], notably ND2 and ND6, for species identification and delimitation in flesh flies.
The heterogeneity of the sequence divergence was examined (Figure S3). The dataset PCGsrRNA showed the highest heterogeneous pairwise sequence divergence, while that of the dataset PCGs was higher than PCGs12. Therefore, the rRNA genes and third-codon positions are the main reasons for such nucleotide heterogeneity, and consequently, excluding the third-codon positions from the datasets would reduce the degree of sequence heterogeneity.

3.5. Phylogenetic Reconstruction

Phylogenetic analyses using the four datasets yielded similar topologies (Figures 5 and S4–S7). The monophyly of Sarcophagidae received maximum support by all reconstructions. All three subfamilies were supported as monophyletic, and their relationship was consistently inferred as (Sarcophaginae, (Paramacronychiinae + Miltogramminae)) (BSs = 95–100; Figures 5 and S4–S7). These results are in agreement with recent phylogenetic studies based on molecular data [2,3,36,49] but conflict with the older hypothesis based on morphology, which supports a sister-group relationship between Paramacronychiinae and Sarcophaginae [4,57].

The genus-level relationships of each subfamily varied significantly across reconstructions based on different matrices. The inclusion of third-codon positions resulted in increased uncertainty of the phylogenetic positions of *A. mihalyii* and *Goniophytus honshuensis* within Paramacronychiinae. When the third-codon positions were excluded, the genus-level relationship within Paramacronychiinae was consistent with the phylogenomic studies [2] and well-supported. However, the pruning of third-codon positions caused instability of relationship within the Sarcophaginae, with either *Blaesoxipha* or *Boettcheria* being a basal branch, compared with (*Blaesoxipha + Boettcheria*) being a sister group to the remaining Sarcophaginae when third-codon positions were included. In contrast, the genus-level relationships of Miltogramminae never reached agreement no matter which dataset was used in this study, and the topologies were generally poorly supported. This may be explained by the higher evolutionary rate of Miltogramminae compared to Paramacronychiinae and Sarcophaginae (Figure 3), and the mitochondrial genome may not be sufficiently phylogenetically informative for resolving older splits in the fast-evolving miltogrammines. In addition, the third-codon positions are generally considered less informative and therefore are often excluded in phylogenetic analyses [58–60], but a recent study suggests third-codon positions to be phylogenetically informative and recommends that they be included [21]. In the present study, the substitutions of third-codon positions of Miltogramminae and Paramacronychiinae were saturated, compared with unsaturated third-codon positions of the Sarcophaginae (Figure S8), which is likely the reason for the different performance of third-codon positions in phylogenetic analyses of Paramacronychiinae and Sarcophaginae. Taken together, our analyses indicate that the mitochondrial genome may be a suboptimal choice for resolving older phylogenetic splits of fast-evolving organisms, but the third-codon positions could be informative in phylogenetic inference.

There are several genera with multiple mitogenomes from GenBank: *Oxysarcodexia, Peckia, Sarcophaga*, and *Wohlfahrtia*. The relationships of species within each of these genera were robust and well-supported in all topologies except for *Sarcophaga*, the largest genus of Sarcophagidae. The relationships within *Sarcophaga* varied depending upon whether the third-codon positions or rRNA genes were used (Figures 5 and S4–S7). Such differences may be a result of fast speciation within *Sarcophaga*, as indicated by the short branches for species of this genus. However, the phylogenetic topology for *Sarcophaga* spp. (Figure 5) largely matches the traditional subgeneric divisions [4].
Figure 5. Maximum likelihood phylogenetic tree for Sarcophagidae based on concatenation of 13 mitochondrial protein-coding genes and 2 rRNA genes. Each node is provided with the support values of Maximum Likelihood bootstrap values (BS) for tree inference using matrix PCGs-rRNA/PCGs12rRNA/PCGs12. Asterisks (*) represent full support, while hyphens (-) indicate that the branch is not recovered.

It is interesting to notice that S. pauciseta (represented by mitogenomic data with accession of NC_053729) always grouped with Blaesoxipha, far from other Sarcophaga. This sequence had been generated as part of a project termed “The complete mitochondrial genome of Sarcophaga rossica” from an unpublished study in NCBI, while “rossica” is the name of a species of Blaesoxipha [4]. In addition, the identification of Sarcophaginae mainly relies on the complicated morphological characters of male terminalia [5], and the general adult morphology of Blaesoxipha is very similar to that of Sarcophaga. Therefore, it is likely that the specimen generating mitogenomic data for S. pauciseta was a misidentified specimen of Blaesoxipha, which highlights the importance of building DNA barcode libraries for insects, such as flesh flies, that are difficult to identify based on morphological characteristics.
4. Conclusions

The current study documents newly sequenced mitogenomes of five species of flesh flies. The mitogenomes examined here are highly conserved in terms of gene content, gene size, gene order, base composition, and codon use of the protein-coding genes, which is consistent with earlier investigations of sarcophagid mitogenomes. Similar to the mitochondrial genomes of other flesh flies, they have a high A + T bias in their nucleotide composition. Among the protein-coding genes, ATP8 appears to have the highest evolutionary rate, while COI has the lowest. Moreover, ND2 and ND6 are shown to exhibit excellent potential as molecular markers for species identification and the construction of phylogenetic relationships. The phylogenetic analysis strongly supports the monophyly of Sarcophagidae and each subfamily, inferring the relationship (Sarcophaginae, (Miltogramminae, Paramacronychiinae)) at the subfamily level, whereas the genus-level phylogeny remains uncertain from the perspective of mitogenomic data. The third-codon positions of the mitochondrial protein-coding genes contribute to the phylogeny construction of Sarcophagidae, particularly within Sarcophaginae.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/insects13080718/s1, Figure S1: GC content of 13 mitochondrial protein-coding genes in Calyptratae. (GC: Guanine and Cytosine). The sequences for species with names in bold were generated in this study, Figure S2: Codon usage analysis of the mitochondrial genomes from three subfamilies (a: Miltogrammatinae, b: Paramacronychiinae and c: Sarcophaginae), Figure S3: Heatmaps showing pairwise heterogeneous sequence divergence of matrices used for phylogenetic analyses. Abbreviations: PCGsrRNA (a), all 13 protein-coding genes and two rRNA; PCGs12 (b), 1st- and 2nd- codon positions of the 13 protein-coding genes; PCGs (c), all 13 protein-coding genes; PCGs12rRNA (d), 1st- and 2nd- codon positions of the 13 protein-coding genes and two rRNA, Figure S4: Maximum likelihood phylogenetic tree for Sarcophagidae based on concatenation of 13 mitochondrial protein-coding genes and two rRNA genes. Each node is provided with the support values of Maximum Likelihood bootstrap values (BS) for tree inference using matrix PCGsrRNA. Asterisks (*) represent full support, Figure S5: Maximum likelihood phylogenetic tree for Sarcophagidae based on concatenation of 13 mitochondrial protein-coding genes. Each node is provided with the support values of Maximum Likelihood bootstrap values (BS) for tree inference using matrix PCGs. Asterisks (*) represent full support, Figure S6: Maximum likelihood phylogenetic tree for Sarcophagidae based on concatenation of 13 mitochondrial protein-coding genes excluding 3rd-codon positions and two rRNA genes. Each node is provided with the support values of Maximum Likelihood bootstrap values (BS) for tree inference using matrix PCGs12. Asterisks (*) represent full support, Figure S7: Maximum likelihood phylogenetic tree for Sarcophagidae based on concatenation of 13 mitochondrial protein-coding genes excluding 3rd-codon positions. Each node is provided with the support values of Maximum Likelihood bootstrap values (BS) for tree inference using matrix PCGs12rRNA. Asterisks (*) represent full support, Figure S8: Nucleotide substitution saturation plots of all the 13 PCGs of three subfamilies based on 1st-, 2nd-, and 3rd- codon positions (a, b, c: Miltogrammatinae; d, e, f: Paramacronychiinae; g, h, i: Sarcophaginae). Transition and transversion are represented by blue and green plots, respectively; Table S1: Collection information of specimens of flesh flies used for DNA extraction in this study, Table S2: The sequence characteristics of mitochondrial genome of flesh flies, Table S3: Nucleotide composition of the mitogenomes of flesh flies.

Author Contributions: Conceptualization, J.S., W.X., T.P., L.Y. and D.Z.; Methodology, J.S., W.X. and L.Y.; Software, J.S. and W.X.; Formal analysis, J.S. and L.Y.; Data curation, J.S., W.X. and X.H.; Writing of original draft, J.S., X.H. and T.P.; Validation, L.Y. and D.Z.; Investigation, D.Z.; Resources, D.Z.; Visualization, L.Y.; Project administration, J.S., W.X., L.Y. and D.Z.; Funding acquisition, D.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by Beijing Forestry University ‘College Student Research and Career-creation Program of Beijing’ (item number: 202010022168), the National Natural Science Foundation of China (32170450, 31872964) and the Beijing Forestry University Outstanding Young Talent Cultivation Project (2019JQ0318).
Institutional Review Board Statement: The collection of samples used in the present study was undertaken with the approval of the School of Ecology and Nature Conservation, Beijing Forestry University.

Informed Consent Statement: Not applicable.

Data Availability Statement: The five newly sequenced mitogenomes were submitted to the GenBank database under the accession numbers of Agria mihaijiyi (ON360967), Blaesoxipha lapidosa (OM640654), Taxigramma karakalensis (ON375459), Wohlfahrtia balassogloi (ON411642), and Wohlfahrtia fellschankoi (ON411643). The associated raw data produced for this study can be found in the NCBI Sequence Read Archive (SRA) under accession numbers PRJNA684426 and PRJNA844010.

Acknowledgments: The authors are grateful to Chao Wang (Chinese Center for Disease Control and Prevention) and Ming Zhang (Institute of Zoology, Chinese Academy of Sciences) for collecting specimens.

Conflicts of Interest: No potential conflict of interest was reported by the authors.

References

1. Pape, T.; Blagoderov, V.; Mostovski, M.B. Order Diptera Linnaeus, 1758. In Animal Biodiversity: An Outline of Higher-Level Classification and Assessment of Taxonomic Richness; Zhang, Z.-Q., Ed.; Magnolia Press: Auckland, New Zealand, 2011; Volume 3148, pp. 222–229.
2. Yan, L.; Buenaventura, E.; Pape, T.; Kutty, S.N.; Bayless, K.M.; Zhang, D. A Phylotranscriptomic Framework for Flesh Fly Evolution (Diptera, Calyptratae, Sarcophagidae). Cladistics 2021, 37, 540–558. [CrossRef]
3. Buenaventura, E.; Szpila, K.; Cassel, B.K.; Wiegmans, B.M.; Pape, T. Anchored Hybrid Enrichment Challenges the Traditional Classification of Flesh Flies (Diptera: Sarcophagidae). Syst. Entomol. 2020, 45, 281–301. [CrossRef]
4. Pape, T. Catalogue of the Sarcophagidae of the World; Associated Publishers: Gainesville, FL, USA, 1996; Volume 8, pp. 1–58.
5. Buenaventura, E.; Pape, T. Phylogeny, Evolution and Male Terminalia Functionality of Sarcophaginae (Diptera: Sarcophagidae). Zool. J. Linn. Soc. 2018, 183, 808–906. [CrossRef]
6. Pawi, M.; Pape, T.; Deja-Sikora, E.; Sikora, M.; Akbarzadeh, K.; Szpila, K. Molecular Phylogeny of Miltothriinae (Diptera: Sarcophagidae): Implications for Classification, Systematics and Evolution of Larval Feeding Strategies. Mol. Phylogenet. Evol. 2017, 116, 49–60. [CrossRef]
7. Pawi, M.; Szpila, K.; Grzypwacz, A.; Pape, T. A Large-Scale Molecular Phylogeny of Flesh Flies (Diptera: Sarcophagidae). Syst. Entomol. 2014, 39, 783–799. [CrossRef]
8. Ren, L.; Shang, Y.; Chen, W.; Meng, F.; Cai, J.; Zhu, G.; Chen, L.; Wang, Y.; Deng, J.; Guo, Y. A Brief Review of Forensically Important Flesh Flies (Diptera: Sarcophagidae). Forensic Sci. Res. 2018, 3, 16–26. [CrossRef] [PubMed]
9. Szpila, K.; Voss, J.G.; Pape, T. A New Dipteran Forensic Indicator in Buried Bodies. Med. Vet. Entomol. 2010, 24, 278–283. [CrossRef] [PubMed]
10. Jordans, K.; Sonet, G.; Richet, R.; Dupont, E.; Braet, Y.; Desmyter, S. Identification of Forensically Important Sarcophaga Species (Diptera: Sarcophagidae) Using the Mitochondrial COI Gene. Int. J. Legal Med. 2013, 127, 491–504. [CrossRef] [PubMed]
11. Hall, M.J.R.; Farkas, R. Traumatic Myiasis of Humans and Animals. In Contributions to a Manual of Palaearctic Diptera; Papp, L., Darvas, B., Eds.; Science Herald: Budapest, Hungary, 2000; Volume 1, pp. 751–768.
12. Szpila, K.; Hall, M.J.R.; Wardhana, A.H.; Pape, T. Morphology of the First Instar Larva of Obligatory Traumatic Myiasis Agents (Diptera: Calliphoridae, Sarcophagidae). Parasitol. Res. 2014, 113, 1629–1640. [CrossRef]
13. Sivinski, J.; Marshall, S.; Petersson, E. Kleptoparasitism and Phoresy in the Diptera. Florida Entomol. 1999, 82, 179–197. [CrossRef]
14. Iyengar, E.V. Kleptoparasitic Interactions throughout the Animal Kingdom and a Re-Evaluation, Based on Participant Mobility, of the Conditions Promoting the Evolution of Kleptoparasitism. Biol. J. Linn. Soc. 2008, 93, 745–762. [CrossRef]
15. Hall, M.J.R.; Wall, R.L.; Stevens, J.R. Traumatic Myiasis: A Neglected Disease in a Changing World. Annu. Rev. Entomol. 2016, 61, 159–176. [CrossRef]
16. Kuhlmann, U. Biology and Predation Rate of the Sarcophagid Fly, Agria mamillata a Predator of European Small Ermine Moths. Int. J. Pest Manag. 1995, 41, 67–73. [CrossRef]
17. Kurahashi, H. The Systematic Position of Nemapoea cicadina Kato, 1943, a Parasite of the Cicada, Tanna japonensis Distant. Jpn. J. Syst. Entomol. 1996, 2, 109–113. [CrossRef]
18. Hall, M.J.R.; Adams, Z.J.O.; Wyatt, N.P.; Testa, J.M.; Edge, W.; Nikolausz, M.; Farkas, R.; Ready, P.D. Morphological and Mitochondrial DNA Characters for Identification and Phylogenetic Analysis of the Myiasis-Causing Flesh Fly Wohlfahrtia magnifica and its Relatives, with a Description of Wohlfahrtia monegrosensis sp. n. Wyatt & Hall. Med. Vet. Entomol. 2009, 23, 59–71. [CrossRef]
19. Hall, M.; Wall, R. Myiasis of Humans and Domestic Animals. Adv. Parasitol. 1995, 35, 257–334. [CrossRef]
20. Cameron, S.L. Insect Mitochondrial Genomics: Implications for Evolution and Phylogeny. Annu. Rev. Entomol. 2014, 59, 95–117. [CrossRef]
49. Yan, L.; Zhang, M.; Gao, Y.; Pape, T.; Zhang, D. First Mitogenome for the Subfamily Miltogramminae (Diptera: Sarcophagidae) and Its Phylogenetic Implications. *Eur. J. Entomol.* 2017, 114, 422–429. [CrossRef]

50. Ren, L.; Zhang, X.; Li, Y.; Shang, Y.; Chen, S.; Wang, S.; Qu, Y.; Cai, J.; Guo, Y. Comparative Analysis of Mitochondrial Genomes among the Subfamily Sarcophagininae (Diptera: Sarcophagidae) and Phylogenetic Implications. *Int. J. Biol. Macromol.* 2020, 161, 214–222. [CrossRef]

51. Ojala, D.; Monroya, J.; Attard, G. TRNA Punctuation Model of RNA Processing in Human Mitochondria. *Nature* 1981, 290, 470–474. [CrossRef]

52. Hebert, P.D.N.; Cywinska, A.; Ball, S.L.; DeWaard, J.R. Biological Identifications through DNA Barcodes. *Proc. R. Soc. B Biol. Sci.* 2003, 270, 313–321. [CrossRef]

53. Demari-Silva, B.; Foster, P.G.; de Oliveira, T.M.P.; Bergo, E.S.; Sanabani, S.S.; Pessôa, R.; Salum, M.A.M. Mitochondrial Genomes and Comparative Analyses of *Culex camposi*, *Culex coronator*, *Culex usquatus* and *Culex usquatissimus* (Diptera:Culicidae), Members of the Coronator Group. *BMC Genom.* 2015, 16, 831. [CrossRef]

54. Du, Y.; Dai, W.; Dietrich, C.H. Mitochondrial Genomic Variation and Phylogenetic Relationships of Three Groups in the Genus Scaphoideus (Hemiptera: Cicadellidae: Deltocephalinae). *Sci. Rep.* 2017, 43, 344–351. [CrossRef]

55. Ma, L.; Liu, F.; Chiba, H.; Yuan, X. The Mitochondrial Genomes of Three Skippers: Insights into the Evolution of the Family Hesperiidae (Lepidoptera). *Genomics* 2020, 112, 432–441. [CrossRef]

56. Verves, Y.G. To the Knowledge of the Subfamilies of the Sarcophagidae (Diptera). *Int. J. Dipterol. Res.* 1998, 9, 243–244.

57. Zhao, Z.; Su, T.; Chesters, D.; Wang, S.; Ho, S.Y.W.; Zhu, C.; Chen, X.; Zhang, C. The Mitochondrial Genome of Elodia flavipalpis Aldrich (Diptera: Tachinidae) and the Evolutionary Timescale of Tachinid Flies. *PLoS ONE* 2013, 8, e61814. [CrossRef] [PubMed]

58. Hassanin, A. Phylogeny of Arthropoda Inferred from Mitochondrial Sequences: Strategies for Limiting the Misleading Effects of Multiple Changes in Pattern and Rates of Substitution. *Mol. Phylogenet. Evol.* 2006, 38, 100–116. [CrossRef] [PubMed]

59. Pons, J.; Ribera, I.; Bertranpetit, J.; Balke, M. Nucleotide Substitution Rates for the Full Set of Mitochondrial Protein-Coding Genes in Coleoptera. *Mol. Phylogenet. Evol.* 2010, 56, 796–807. [CrossRef] [PubMed]