Characterization of the complete mitochondrial genome of the Sunda stink-badger (*Mydaus javanensis*) from the island of Borneo (#83689)

First revision

Guidance from your Editor

Please submit by **2 Aug 2024** for the benefit of the authors (and your token reward) .

Structure and Criteria
Please read the 'Structure and Criteria' page for guidance.

Custom checks
Make sure you include the custom checks shown below, in your review.

Author notes
Have you read the author notes on the guidance page?

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

If this article is published your review will be made public. You can choose whether to sign your review. If uploading a PDF please remove any identifiable information (if you want to remain anonymous).

Files

Download and review all files from the materials page.

1 Tracked changes manuscript(s)
1 Rebuttal letter(s)
9 Figure file(s)
9 Table file(s)
1 Other file(s)

Custom checks

DNA data checks

- Have you checked the authors data deposition statement?
- Can you access the deposited data?
- Has the data been deposited correctly?
- Is the deposition information noted in the manuscript?

Vertebrate animal usage checks

- Have you checked the authors ethical approval statement?
- Were the experiments necessary and ethical?

Have you checked our animal research policies?

Field study

Have you checked the authors field study permits?

Are the field study permits appropriate?

For assistance email peer.review@peerj.com

Structure and Criteria

2

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- **Impact and novelty is not assessed.** Meaningful replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Characterization of the complete mitochondrial genome of the Sunda stink-badger (*Mydaus javanensis*) from the island of Borneo

Vijay Kumar Subbiah [Corresp., 1], **Chrishen Gomez** [1, 2], **Dexter Miller Robben** [1], **Ranjita Subramaniam** [1], **Andrew J Hearn** [2]

[1] Biotechnology Research Insitute, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia

[2] Wildlife Conservation Research Unit, Department of Biology, University of Oxford, Oxfordshire, United Kingdom

Corresponding Author: Vijay Kumar Subbiah
Email address: vijay@ums.edu.my

Background: The Mephitidae is a family of skunks and stink-badgers that includes 12 extant species in four genera namely, *Mydaus, Conepatus*, *Mephitis* and *Spilogale*. *Mydaus* is the only genus within Mephitidae found outside the American continent, with its distribution limited to the islands of Borneo, Indonesia and Philippines. There are two extant species of *Mydaus* i.e. *javanensis* and *marchei.* Currently, complete mitogenomes are unavailable for either species. Here, we present the characterization of the first complete mitogenome for the Sunda stink-badger (*Mydaus javanensis*) from the island of Borneo. **Methods:** Muscle tissue was obtained and the DNA was sequenced using a combination of Illumina Barcode Tagged Sequence (BTSeq) and Sanger sequencing techniques. The genome was annotated with MITOS and manually checked for accuracy. A circular map of the mitogenome was constructed with Proksee. Relative synonymous codon usage (RSCU) and codon frequency were calculated using MEGA-X. The protein coding genes (PCGs) were aligned with reference sequences from GenBank and used for the construction of phylogenetic trees (ML and BI). Additionally, due to the lack of available complete genomes in public databases, we constructed another tree with the *cyt b* gene. **Results:** The complete circular mitogenome was 16,391 base pairs in length. It comprises the typical 13 protein-coding genes, 22 tRNAs, two ribosomal RNA genes, one control region (CR) and an L-strand replication origin (O_L). The G+C content was 38.1% with a clear bias towards A and T nucleotides. Of the 13 PGCs, only ND6 was positioned in the reverse direction, along with five other tRNAs. Five PCGs had incomplete stop codons and rely on post-transcriptional polyadenylation (TAA) for termination. Based on the codon count, Leucine was the most common amino acid (589), followed by Threonine (332) and Isoleucine (325). The ML and BI phylogenetic trees, based on concatenated PCGs and *cyt b* gene, respectively, correctly clustered the species with other members of the Mephitidae family but were unique enough to set it apart from *Conepatus*, *Mephitis* and *Spilogale*. The

result confirms *Mydaus* as a member of the mephitids and the mitogenome will be useful for evolutionary analysis and conservation of the species.

Characterization of the complete mitochondrial genome of the Sunda stink-badger (*Mydaus javanensis*) from the island of Borneo

Vijay Kumar Subbiah[a], Chrishen Gomez[a,b], Dexter Miller Robben[a], Ranjita Subramaniam[a] and Andrew J Hearn[b*]

[a]Biotechnology Research Institute, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia;

[b]Wildlife Conservation Research Unit, Department of Biology, University of Oxford, UK.

* Corresponding Author:
Vijay Kumar Subbiah
[a]Biotechnology Research Institute,
Universiti Malaysia Sabah,
88400 Kota Kinabalu,
Sabah,
Malaysia.

Email address: vijay@ums.edu.my

Abstract

Background: The Mephitidae is a family of skunks and stink-badgers that includes 12 extant species in four genera namely, *Mydaus, Conepatus*, *Mephitis* and *Spilogale*. *Mydaus* is the only genus within Mephitidae found outside the American continent, with its distribution limited to the islands of Borneo, Indonesia and Philippines. There are two extant species of *Mydaus* i.e. *javanensis* and *marchei.* Currently, complete mitogenomes are unavailable for either species. Here, we present the characterization of the first complete mitogenome for the Sunda stink-badger (*Mydaus javanensis*) from the island of Borneo.

Methods: Muscle tissue was obtained and the DNA was sequenced using a combination of Illumina Barcode Tagged Sequence (BTSeq) and Sanger sequencing techniques. The genome was annotated with MITOS and manually checked for accuracy. A circular map of the mitogenome was constructed with Proksee. Relative synonymous codon usage (RSCU) and codon frequency were calculated using MEGA-X. The protein coding genes (PCGs) were aligned with reference sequences from GenBank and used for the construction of phylogenetic trees (ML and BI). Additionally, due to the lack of available complete genomes in public databases, we constructed another tree with the *cyt b* gene.

Results: The complete circular mitogenome was 16,391 base pairs in length. It comprises the typical 13 protein-coding genes, 22 tRNAs, two ribosomal RNA genes, one control region (CR) and an L-strand replication origin (O_L). The G+C content was 38.1% with a clear bias towards A and T nucleotides. Of the 13 PGCs, only ND6 was positioned in the reverse direction, along with five other tRNAs. Five PCGs had incomplete stop codons and rely on post-transcriptional polyadenylation (TAA) for termination. Based on the codon count, Leucine was the most common amino acid (589), followed by Threonine (332) and Isoleucine (325). The ML and BI phylogenetic trees, based on concatenated PCGs and *cyt b* gene, respectively, correctly clustered the species with other members of the Mephitidae family but were unique enough to set it apart from *Conepatus*, *Mephitis* and *Spilogale*. The result confirms *Mydaus* as a member of the mephitids and the mitogenome will be useful for evolutionary analysis and conservation of the species.

Keywords: Stink-badger; Mitogenome; *Mydaus javanensis*; Mephitids; Borneo; Skunk

INTRODUCTION

Stink-badgers and skunks are grouped under the family Mephitidae, which includes 12 extant species placed in four genera: *Mydaus* (stink-badgers), *Conepatus* (hog-nosed skunks), *Mephitis* (hooded and striped skunks), and *Spilogale* (spotted skunks). Three other genera under Mephitidae (*Brachyprotoma*, *Palaeomephitis,* and *Promephitis*) are extinct and known only through fossil records. The stink-badger *Mydaus* is the only one within Mephitidae found outside the American continent. Its resemblance to badgers led early authors to classify the species under Mustelidae (subfamily: Melinae), but recent molecular evidence has led to its reclassification as members of Mephitidae (Dragoo & Honeycutt 1997; Hwang and Lariviere 2003). To date, the genus includes two species: *Mydaus javanensis* and its only known sister taxon, *Mydaus marchei*.

The Sunda stink-badger (*Mydaus javanensis*), commonly known as the Malay Badger or Teledu, is one of Southeast Asia's least-studied carnivores. Its biology and natural history remain poorly understood. It is thought to be tolerant of anthropogenic disturbance and inhabits a wide variety of habitat types, including primary and disturbed forests, open areas adjacent to forests, and oil palm plantations (Md-Zain *et al*., 2019). Within Borneo, this species is recorded frequently in the Malaysian state of Sabah, northern Borneo (Samejima *et al*., 2016; Wong *et al*., 2017), but much less frequently in South Kalimantan (Higashide *et al*., 2018) and exhibits a much localized distribution in northern Sarawak (Giman and Jukie, 2012). The driver of this patchy distribution across Borneo is presently unknown (Samejima *et al*., 2016; Wong *et al*., 2017).

Previous documentation indicates that access to mitochondrial genetic data from members of Mephitidae has been crucial for accurately defining species boundaries, identifying unrecognized species diversity within a geographic region, tracing complex evolutionary histories (such as secondary contacts between insular populations), understanding the influence of factors like climate change on phylogeographic diversification, and determining the timing and methods of specific geographic colonization. Such information plays a crucial role in guiding conservation decisions for the mephitids (McDonough *et al*., 2022; Bolas *et al*., 2022).

The absence of a complete mitochondrial genome sequence has been a drawback for *Mydaus*, as existing research entails limited choices of gene markers or the usage of partial genes (Dragoo & Honeycutt 1997; Md-Zain *et al*., 2021). A complete mitochondrial genome for *Mydaus* will open up avenues for future investigations to ascertain true mitochondrial lineages. A profound understanding of mitochondrial sequence characterization can shed light on suitable regions to use as genetic markers. Additionally, a complete mitogenome may offer accurate signals for phylogenetic reconstruction compared to gene fragments (Lan *et al*., 2024). In this context, having a complete reference mitochondrial genome representing a genus can aid in the sequencing and assembly of additional species from the taxa.

Overall, the complete mitogenome sequence of *M. javanensis* will provide a wealth of information on its evolutionary history, genetic diversity, population dynamics, and relationships

with other members of the Mephitidae. Thus, in the current study, we aim to determine and characterize the complete mitochondrial sequence of *M. javanensis*. This effort will not only add new information but also aid in conservation efforts and management strategies for this understudied member of the Mephitidae family.

MATERIALS AND METHODS

Sample collection and identification

Muscle tissues were collected from an adult male Sunda stink-badger (sample number MJ19) found as a road-killed animal during a routine field sampling trip in the Tawau region of Sabah, Malaysian Borneo (4º19'54.6" N 117º 52'03.9" E). The specimen was placed in ice during field sampling and subsequently stored in a -20 °C freezer until further use. The sample was collected as part of a study on the ecology and conservation of Bornean carnivores, under an access license permit granted by the Sabah Biodiversity Centre (JKM/MBS.1000-2/2 JLD.12(48)) and under ethical review by the University of Oxford (Ref. No. APA/1/5/ZOO/NASPA/WildCRU/BorneanCarnivores).

DNA extraction, PCR amplification and sequencing

DNA was isolated from approximately 25mg tissue using the DNeasy Blood and Tissue DNA Extraction Kit (Qiagen, USA) according to the manufactures protocol. After quality control using gel electrophoresis and UV spectrophotometer, the DNA sample was PCR amplified with newly designed primers (Table 1) to capture the entire ~16kb mitochondrial DNA following a similar procedure developed by Deiner *et al*. (2017). For this purpose, we utilized the high-fidelity PrimeSTAR® GXL polymerase (Takara Bio, Japan), specifically designed for long-range PCR amplification for this purpose. The PCR was carried out in a total volume of 25 µL containing 30 ng of genomic DNA, 0.4 U of PrimeSTAR® GXL polymerase, 1x PrimerSTAR GXL buffer, 0.5 mm dNTPs (PrimerSTAR), and 10 pmol of primers Leo16SLRpcr_F and Leo16SLRpcr_R. The PCR amplification was performed as follows: pre-denaturation at 98 °C (10 sec), followed by 40 cycles of denaturation at 98 °C (10 sec), annealing at 60 °C (15 sec) and extension at 68 °C (14 min). A final extension step at 68 °C for 5 min was included. The PCR product was then electrophoresed on a 0.8 % agarose gel with 1x TBE buffer. The amplicon was excised from the gel and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega, USA). The sample was then sequenced using the Barcode Tagged Sequencing (BTSeq^{TM}) approach on an Illumina platform. Briefly, barcoded adapters are used to tag DNA fragments, which were then subsequently sequenced on an Illumina platform. Celemics' exclusive bioinformatics pipeline then organizes the sequencing reads based on molecular barcodes and aggregates them to rectify NGS errors, resulting in the generation of a complete

DNA sequence. The sequencing service was provided by Celemics Inc (http://www.celemics.com).

In order to capture the regions flanking the primer binding sites of the circular mitogenome, we designed another set of primers (Table 1) to amplify a 688 bp region. This PCR was carried out as described above but with 15 pmol of primers MJ19F and MJ19R instead. The PCR amplification was as follows: pre-denaturation at 95 °C (5 min), followed by 40 cycles of denaturation at 95 °C (30 sec), annealing at 55 °C (30 sec) and extension at 72 °C (45 sec). A final extension step at 72 °C for 10 min was included. The PCR product was electrophoresed on a 1.5% agarose gel, purified and subsequently sequenced bi-directionally using BigDye Terminator v3.1 on an ABI3130 Sequencer.

Mitogenome annotation and analysis

The two high-quality reads from BTSeq and Sanger sequencing were manually assembled to form complete contiguous circular reads. The mitogenome was annotated using MITOS webserver (http://mitos2.bioinf.uni-leipzig.de; Donath *et al*. 2019). To ensure accuracy, the annotation, intergenic spacers, and overlapping regions between genes were manually checked, counted and compared with complete and near complete mitogenomes of other related taxa from NCBI. All boundaries and secondary structures of tRNA gene were crossed-checked with tRNAscan-SE v2.0 (http://lowelab.ucsc.edu/tRNAscan-SE; Chan & Lowe, 2019) with the parameters: source = “Mito/Chloromast” and genetic code = “Vertebrate Mito and ARWEN v1.2 (Laslett & Canbäck, 2008), under default settings. A circular map of the mitogenome with all its respective features was drawn using the Proksee online tool (https://proksee.ca).

Base composition and relative synonymous codon usage (RSCU) were analyzed using MEGA-X (Kumar et al., 2018). Strand asymmetry was calculated using the formulas by Perna & Kocher (1995) i.e. AT skew = (A - T)/(A + T) and, GC skew = (G - C)/(G + C).

Characterization of the control region (CR)

The control region (CR) sequence of *M. javanensis* was mined from its complete mitogenome, which was sequenced in this study. The organization of *M. javanensis* was compared with those of other mephitids, whose CR regions were retrieved from their respective complete mitogenomes downloaded from NCBI GenBank. Incomplete sequences with gaps were discarded, prior to further analysis. The remaining sequences were examined for termination of the displacement loop (D-loop) motif, termination-associated sequences (TAS-A), and putative conserved sequence blocks, according to previous reports for *Lutra lutra* (Eurasian otter), *Conepatus chinga*, and *Conepatus leuconotus leuconotus* (Ketmaier & Bernardini, 2005). The

alignment of these structural features was performed using MUSCLE in MEGA X, version 10.2.6 (Kumar *et al*., 2018). The repeats in the CR region among the available mephitids were identified using the Tandem Repeat Finder program (http://tandem.bu.edu/trf/trf.html) (Benson, 1999), under default settings.

Phylogenetic analysis

In order to infer the phylogenetic relations of *M. javanensis* and other mephitids, concatenated nucleotide sequences were generated based upon 13 PCGs of the mitogenome. We also included members from Mustelidae, Procyonidae, Ailuridae and from the suborder Caniformia for comparison purposes. *Neofelis nebulosa* (Suborder: Feliformia) was treated as an outgroup. The sequences were downloaded from the Genbank. Complete mitogenomes were not available for *Conepatus humboldtii*, *Mephitis macroura* and *Mydaus marchei* and thus, excluded from the tree.

PartitionFinder 2.1.1 (Lanfear *et al*., 2017) was used to select the best substitution models and partition schemes (Supplementary Data S1) with "greedy" algorithm and Bayesian information criterion (BIC), to be used in the subsequent phylogenetic analyses. Maximum Likelihood (ML) and Bayesian Inference (BI) approaches were applied for these analyses. The ML analysis was performed in IQ-TREE version 1.6.12 (Nguyen *et al*., 2015) with the node reliability assessed with 1,000 replicates of ultrafast likelihood bootstrap (Minh *et al*., 2013). The BI analysis was conducted with MrBayes on XSEDE v3.2.7a (Ronquist *et al*., 2012) available through the CIPRES Science Gateway (https://www.phylo.org/) (Miller, Pfeiffer & Schwartz, 2010). The Markov chain Monte Carlo (MCMC) runs were conducted for 10,000,000 generations and the trees were sampled every 1,000 generations with a burn-in of 25%. The software Tracer v1.7.2 (Rambaut *et al*., 2018) was employed to assess the parameters (effective sampling size for all parameters > 200).

In addition, we conducted phylogenetic analyses (ML and BI) using only *cyt b* gene sequences due to the scarcity of whole mitogenomes for the mephitids. Our analysis encompassed all accessible *cyt b* gene sequences from GenBank for members of the Mephitidae family. The first 37 nucleotide bases of the gene were trimmed for all species, considering the availability of partial gene for *Spilogale putorius putorius* and *Spilogale putorius ambarvalis*. The ML analysis was performed in IQ-tree version 1.6.12 (Nguyen *et al*., 2015) based on the best-substitution model (TIM2+F+I+G4) selected by ModelFinder (Kalyaanamoorthy *et al*. 2017) in the IQ-TREE package with 1,000 ultrafast bootstrap replicates (Minh *et al*., 2013). The BI analysis was executed with Mrbayes on XSEDE v3.2.7a (Ronquist *et al*., 2012) where the best-fit substitution model (TIM2+I+G) was determined via Jmodeltest2 (Darriba *et al*., 2012) and these were available through CIPRES Science Gateway (https://www.phylo.org/) (Miller, Pfeiffer & Schwartz, 2010). The Markov chain Monte Carlo (MCMC) runs were performed for

10,000,000 generations and the trees were sampled every 1,000 generations with a burn-in of 25%. Tracer v1.7.2 (Rambaut *et al*., 2018) was used to assess the parameters (effective sampling size for all parameters > 200). The resulting trees were visualized using Figtree v1.4.4.

RESULTS

Mitogenome organization and composition

The complete mitogenome of *M. javanensis* sample MJ19 was 16,391 bp and the GenBank accession number is OP442081. It was sequenced at a high quality at about 1100× coverage. The size of the mitogenome was within the range of the complete mitogenomes from other mephitids (Table 2). It comprises of the typical 13 protein-coding, 22 tRNAs, two ribosomal RNA genes, one control region (CR) and an L-strand replication origin (O_L). The *ND6* gene, along with eight other tRNAs (*trnQ*, *trnA*, *trnN*, *trnC*, *trnY*, *trnS2*, *trnE* and *trnP*), was positioned on the reverse direction (Fig. 1). All other genes and miscellaneous regions were in the forward direction. The *trnL* and *trnS* were made up of two major codons i.e. *trnL1 (UUR)* and *trnL2 (CUN)* and *trnS1 (AGY)* and *trnS2 (UCN)*, respectively. A total of nine pairs of genes (or regions) overlapped with one another, with overlaps ranging from 1 to 43 nucleotides (Table 3). We noticed that five PCGs had incomplete stop codons and likely rely on post-transcriptional polyadenylation (TAA) for termination.

With regards to the base composition, the mitogenome was skewed with a clear bias towards an A+T content of 62.9%, while G+C was 38.1% (Table 4). Composition analysis revealed that the mitogenome exhibited a positive AT (0.095) and a negative GC skew (-0.307) as a whole, as well as in the 13 PCGs (AT skew: 0.045; GC skew: -0.337), 2 rRNAs (AT skew: 0.201; GC skew: -0.099) and the control region (AT skew: 0.039; GC skew -0.253). However, for the tRNA genes, both the AT and GC skews were positive, at 0.031 and 0.073, respectively (Table 4).

Protein coding genes (PCGs) and codon usage

The mitogenome was comprised of the typical 13 PCGs found in mammals. The concatenated lengths of the 13 PCGs were 11,424 bp and a total of 3,808 codons were involved in protein translation. Based on the codon count, Leucine was the most common amino acid (589), followed by Threonine (332) and Isoleucine (325). It was interesting to note that the RSCU indicated that degenerate codons were biased towards using more A and C at the third codon compared to G and U (Fig. 2).

Transfer RNA and ribosomal RNA genes

The *M. javanensis* mitogenome contained the typical 22 tRNA genes and their lengths ranged from 59 (*trnS1*) to 75 (*trnL2*). All tRNA exhibited the typical cloverleaf secondary

structure, with the exception of *trnS1*^{AGY} which lacks the dihydrouridine loop (Fig. 3). There was a total of 38 mismatches (U-G, U-U, C-A, A-A, A-G and C-U) with the U-G (74%) being the most common.

In addition, the two rRNA genes (*rrnL* and *rrnS*) were highly conserved across the mephitids. The putative lengths of *rrnL* and *rrnS* were 1,572 and 957 bp, respectively, and both ha a positive AT skew and a negative GC skew (Tables 3 and 4).

Control Region (CR)

At position 162 of the *M. javanensis*'s control region, we identified a motif homologous to the termination-associated sequences (TAS-A). The motif of the D-loop termination of the CR (GCCCC) was identified few nucleotides upstream of TAS-A. In addition, all the eight putative conserved sequence blocks (CSB1-3 and B-F) were identified within the CR region. A single region with tandem repeats, referred to as RS3, was discovered in between CSB-1 and CSB-2. A schematic diagram illustrating the organization of the control region of *M. javanensis* is shown in Fig. 4.

A comparative analysis of the control region was conducted according to previous reports on Eurasian otter and other mephitids (Fig. 5). The analysis reveals the presence of fundamental structures, such as the D-loop termination, TAS-A, and multiple conserved sequence blocks, in all the mephitids. The identification and comparison of the tandem repeats among mephitids were also conducted, revealing differences in the distribution of the repeat regions (Table 5). Similar to *M. javanensis*, mephitids such as *Mephitis mephitis*, *Conepatus chinga*, and *Conepatus leuconotus leuconotus* display a single repeat region located in the right domain, between CSB-1 and CSB-2. Conversely, *Spilogale angustifrons* showed the presence of a repeat region in the left domain. In contrast, *Spilogale putorius* exhibited two repeat regions, one in the right domain and the other in the left domain. Interestingly, the Tandem Repeat Finder did not detect any repeats in either of the subspecies of *Spilogale gracilis*. Both *Spilogale gracilis gracilis* and *Spilogale gracilis leucoparia* had the smallest control region sizes, with 819 bp and 853 bp, respectively (Table 5). Overall, the mitochondrial control region maintains conservation in its basic structures, with differences appearing in the tandem repeat regions.

Phylogenetic analysis

Phylogenetic trees of ML (Fig. 6) and BI (Fig. 7) analyses were built on concatenated sequences of 13 PCGs from 30 species, with 13 members representing the Mephitidae. Overall, both the trees were highly congruent and received high bootstrap support for the majority of branches, accurately clustering the species within its specific genera and family. Both the topologies confirmed the monophyly of the family Mephitidae which included the members representing *Spilogales*, *Mephitis*, *Conepatus* and *Mydaus*. *M. javanensis* MJ19 (Borneo, Malaysia) was placed into this monophyletic group together with the other member from Java, Indonesia, with high nodal supports from both the trees (boostrap support values (BS)=100 and

Bayesian posterior probability (BPP) was equal to 1.00). Meanwhile, we have also observed that Mustelidae form sister clade with Procyonidae.

Meanwhile, ML (Fig. 8) and BI (Fig. 9) trees constructed based on *cyt b* gene included additional species within Mephitidae. The resulting trees generated the same topology and BI analysis provided more resolution with strong supports than ML analysis. The trees correctly grouped the Javanese and Bornean Sunda stink-badger with high bootstrap values (BS=100).

DISCUSSION

The mitogenome of the Sunda stink-skunk was sequenced and characterized. This is the first publicly available complete mitogenome for the species. The information provided here has added new information to the relatively understudied members of Mephitidae family. It is important to note that we still do not have mitogenomes for another three mephitids (*Conepatus humboldtii*, *Mephitis macroura* and *Mydaus marchei*), while some mitogenomes (*C. leuconotus, C. semistriatus*, *S. pygmaea*) were incomplete and contained stretches of Ns in their sequences (Table 2). The primers used in the study managed to amplify the entire mitochondrial DNA in a single PCR following the approach originally used by Deiner *et al*. (2017) and Kato-Unoki *et al*. (2020). Targeting the entire mitogenome in a single PCR is preferable as it prevents the amplification of nuclear encoded pseudo mitochondrial genes and avoids the misalignment of gene order during assembly and annotation (Parr *et al*., 2006; Montaña-Lozano *et al*., 2022). However, one drawback of this approach is that DNA sequencing is unable to capture sequences at the primer binding sites, hence, a second set of primers were needed to target this region. Our primers were based on conserved regions and may be useful to other researchers who are working on mephitids or other closely related species. In addition, these primers could be used to fill up the gaps in the partial mitogenomes mentioned in Table 2.

The structure of the *Mydaus javanensis* MJ19 mitogenome was similar with other vertebrates with the typical 13 protein-coding genes, 22 tRNAs, two ribosomal RNA genes, one control region (CR) and an L-strand replication origin (O_L) (Boore, 1999; Pereira, 2000; Montaña-Lozano *et al*., 2022). These components make up the mammalian mitochondrial oxidative phosphorylation (OXPHOS) system (Signes *et al*., 2018; Shokolenko & Alexeyev, 2022). Overlapping among 10 gene pairs was observed and such occurrence has been proposed to extend genetic information within the constraints of limited genome size (Sun *et al*., 2020).

The AT and GC skew of *M. javanensis* was similar to other members of the Mephitidae and to that reported previously in Mustelidae (Skorupski, 2022). It is interesting to note that the bias toward A and T and against G and C is a common feature in metazoan mitogenomes. This naturally leads to a subsequent bias in the corresponding encoded amino acids as seen in the codon usage. However, when compared to members of Proconidae, Ailuridae, Canidae and

Ursidae, we noticed a smaller GC skew (an average of -0.268) in the species (Skorupski, 2022), indicating smaller disproportionate of G bases compared to Cs. This is reflected that in *M. javanensis*, the degenerate codons were biased to use more A and C at the third codon.

With regards to the 22 tRNA genes in *M. javanesis*, we observed a typical and conserved arrangement as found in most vertebrates (Watanabe *et al*., 2014). One unique feature is the lack of the dihydrouridine (DHU) loop in $trnS1^{AGY}$ which is commonly observed in all vertebrates (Pereira, 2000). The DHU loop is important as it is involved in the aminoacylation of the tRNA molecule. The loop functions as a recognition site for aminoacyl-tRNA synthetase (Watanabe *et al*., 2014). In contrast, $trnS2^{UCN}$ has the complete typical cloverleaf pattern features of tRNAs including the DHU loop. While both the $trnS1^{AGY}$ and $trnS2^{UCN}$ are capable of translation on the ribosome, Hanada *et al*. (2001) has shown that the lack of DHU loop in $trnS1^{AGY}$ results inconsiderably lower translational activity. This explains why $trnS2^{UCN}$ is the preferred tRNA for Serine as observed by the codon count and RSCU distribution in *M. javanensis* (Figure 2).

In regard to the analysis of the control region (CR) organization in *M. javanensis*, we discovered the presence of ~~the~~ fundamental conserved structures similar to those previously identified in Eurasian otter (a mustelid) as well as other mephitids, such as the D-loop termination motif, TAS-A and eight conserved sequence blocks (Ketmaier & Bernardini, 2005). Additionally, the CR of mammals known to have two potential locations with tandem repeats: one in the left and the other in the right domains of the gene, respectively (Wilkinson et al. 1997). In *M. javanensis* we found a single repeat region (RS3) which is placed between CSB-1 and CSB-2, similar to various other skunks (Ketmaier & Bernardini, 2005). Notably, this location has only been described in mammals (Ketmaier & Bernardini, 2005). As such, CSB-1 and CSB-2 can serve as potential primer binding sites for future amplifications of RS3 regions, especially for the species under *Mydaus*.

Additionally, a comparative analysis of CR structural organization among mephitids showed that the CR region is well structured with the central region being highly conserved and tandemly repeated sequences occurring only within the two peripheral domains. These peripheral domains are rapidly evolving regions characterized by a high rate of nucleotide substitutions and variations in the copy number of tandem repeats. This variability in the number of tandemly repeated sequences is considered as the pivotal source of mitochondrial DNA length variation in animals (Brown *et al*. 1986). Consistent with this, we found that species lacking the repeat regions had the smallest overall control regions sizes (Table 5). Species within the Mephitidae family exhibited diverse distributions of repeat regions. Some had none, some had two, and those with a single repeat region had it located in either the right or left domains. Interestingly, the distribution of repeat regions was found to be species-specific. In future studies, additional species within the genus *Mydaus* can be investigated to gain further insights into the distribution and evolutionary patterns of tandem repeats in the control region (CR).

The ML and BI trees based on the concatenated sequences of 13 Protein Coding Genes (PCGs) correctly grouped the 9 (out of 12) extant species of mephitids in four 4 genera of the family Mephitidae. *M. javanensis* was accurately grouped with the mephitids and not with the mustelids. The genus *Mydaus* was initially placed as a member of Mustelidae till the revision using DNA studies revealed otherwise that the skunks, along with stink badgers (*Mydaus* spp.) belong to a separate family (Mephitidae) which is highly divergent from the mustelids (Koepfli *et al*., 2017). The phylogeny using the complete set of core protein coding genes of the mitochondrial genome conducted in this study, is in concordance to this finding. Additionally, both the ML and BI trees reflected a common observance of Mustelidae forming sister clade with Procyonidae. The sister-group relationship between these two families is congruent to previous documentations, supported by morphological characters and molecular data (Sato *et al*., 2003).

To utilize the sequences available in GenBank for the mephitids, we used the *cyt b* gene sequences to perform ML and BI analyses. In both topologies, sample *M. javanesis* MJ19, from northern Borneo (Sabah, Malaysia) was grouped together with the Javanese sample from Indonesia with high support values (BS=100 and BPP=1.0). However, we noticed 27 nucleotide variations between the two individuals in the partial *cyt b* gene (the first 37 nucleotide bases were trimmed to standardize the lengths of all sequences under study) (Supplementary Data S2 and S3). All nucleotide variants at the third codon base were degenerate and did not result in changes in the amino acid. However, when the nucleotide variants were in the first codon base, it invariably led to seven non-synonymous amino acid changes (Supplementary Data S4). These results indicate that even within a conserved gene region like the *cyt b*, there are polymorphisms that can differentiate the Sunda stink-badgers from Java and Borneo. It would be worthwhile to sequence the complete mitogenomes from the Javanese species and surrounding Southeast Asian populations to determine its genetic diversity, enhancing our understanding of the evolutionary patterns of *Mydaus*.

CONCLUSION

In the present study, we sequenced the first mitogenome of the Sunda stink-skunk (*M. javanensis*) and presented its structures and characteristics. We observed a consistency in the genome size (16,391 bp) across metazoan mitogenomes. There were nine genes which were found to overlap each other in the mitogenome. Some PCGs lacked complete stop codons, possibly functioning through post-transcriptional polyadenylation for termination. *M. javanensis* showed the typical bias towards A and T which is commonly observed in metazoan mitogenomes.

The mitogenome of *M. javanensis* also featured some unique characteristics. Although the 22 tRNA genes had conserved arrangements, the DHU was absent in the $trnS1^{AGY}$ tRNA

gene. This may lead to lower translational activity, potentially explaining why *trnS2*^{UCN} is preferred for Serine over *trnS1*^{AGY}. Our analysis also shed new light on the organization of the mitochondrial control region (CR) in *M. javanensis* as compared to other mephitids, showing species-specific variation in the presence and distribution of tandem repeats within the CR. The phylogenetic relationships constructed by both ML and BI methods, based on concatenated sequences of 13 PCGs as well as *cyt b* gene, were consistent and supported the monophyly of Mephitidae. The phylogenetic trees clearly placed *M. javanensis* with *Conepatus*, *Mephitis*, and *Spilogale*, confirming its position within the extant species of Mephitidae. The information provided here has added new information to the relatively understudied members of Mephitidae family.

ACKNOWLEDGEMENTS

We thank the Sabah Biodiversity Centre and Sabah Forestry Department for permission to conduct research on Bornean carnivores in Sabah.

REFERENCES

Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* **27(2)**:573-580. DOI: 10.1093/nar/27.2.573

Bolas EC, Quinn CB, Van Vuren DH, Lee A, Vanderzwan SL, Floyd CH, Jones KL, Shaskey L, Sacks BN. 2022. Pattern and timing of mitochondrial divergence of island spotted skunks on the California Channel Islands. *Journal of Mammalogy* **103(2)**:231-242. DOI: 10.1093/jmammal/gyac008

Boore JL. 1999. Animal mitochondrial genomes. *Nucleic Acids Research* **27 (8)**:1767–1780. DOI: 10.1093/nar/27.8.1767.

Brown GG, Gadaleta G, Pepe G, Saccone C, SbisA E. 1986. Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. *Journal of molecular biology* **192(3)**:503-511. DOI: 10.1016/0022-2836(86)90272-x

Chan PP, Lowe TM. 2019. tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. *Methods in Molecular Biology* **1962**:1-14. DOI: 10.1007/978-1-4939-9173-0_1.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772. DOI: 10.1038/nmeth.2109

Deiner K, Renshaw MA, Li Y, Olds BP, Lodge DM, Pfrender ME. 2017. Long-range PCR allows sequencing of mitochondrial genomes from environmental DNA. *Methods in Ecology and Evolution* **8**:1888–1898. DOI: 10.1111/2041-210X.12836.

Donath A, Jühling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, Middendorf M, Bernt M. 2019. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. *Nucleic Acids Research* **47(20)**:10543–10552. DOI: 10.1093/nar/gkz833.

Dragoo JW, Honeycutt RL. 1997. Systematics of Mustelid-Like Carnivores. *Journal of Mammalogy* **78(2)**:426–443. DOI: 10.2307/1382896.

Giman B, Jukie A. 2012. Distribution of Sunda stink-badger Mydaus javanensis in Sarawak, Malaysia. *Small Carnivore Conservation* **46**:12–16.

Higashide D, Tanaka K, Nakama E, Faudi H, Satriadi T, Aryadi M. 2018. Camera-trap records of Sunda Stink-badger Mydaus javanensis and other small carnivores in South Kalimantan, Indonesia. *Small Carnivore Conservation* **56**:54–59.

Hwang YT, Larivière S. 2003. *Mydaus javanensis*. *Mammalian Species* **723**:1-3. DOI: 10.1644/723.

Kato-Unoki Y, Umemura K, Tashiro K. 2020. Fingerprinting of hatchery haplotypes and acquisition of genetic information by whole-mitogenome sequencing of masu salmon,

458 *Oncorhynchus masou masou*, in the Kase River system, Japan. *PLoS ONE* **15(11)**:e0240823.
459 DOI: 10.1371/journal.pone.0240823.
460
461 Ketmaier V, Bernardini, C. 2005. Structure of the mitochondrial control region of the Eurasian
462 otter (*Lutra lutra*; Carnivora, Mustelidae): patterns of genetic heterogeneity and implications for
463 conservation of the species in Italy. *Journal of Heredity* **96(4)**:318-328. DOI:
464 10.1093/jhered/esi037
465
466 Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS. 2017. ModelFinder: fast
467 model selection for accurate phylogenetic estimates. *Nature methods* **14(6)**:587-589. DOI:
468 10.1038/nmeth.4285
469
470 Koepfli KP, Dragoo JW, Wang X. (2017). The evolutionary history and molecular systematics of
471 the Musteloidea. *Biology and Conservation of Musteloids*. Oxford University Press, Oxford, 75-
472 91. DOI: 10.1093/oso/9780198759805.003.0002
473
474 Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary
475 genetics analysis across computing platforms. *Molecular Biology and Evolution* **35(6)**:1547–
476 1549. DOI: 10.1093/molbev/msy096.
477
478 Lan G, Yu J, Liu J, Zhang Y, Ma R, Zhou Y, Zhu B, Wei W, Liu J, Qi G. 2024. Complete
479 Mitochondrial Genome and Phylogenetic Analysis of *Tarsiger indicus* (Aves: Passeriformes:
480 Muscicapidae). *Genes* **15(1)**:90. DOI: 10.3390/genes15010090
481
482 Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017. PartitionFinder 2: new
483 methods for selecting partitioned models of evolution for molecular and morphological
484 phylogenetic analyses. *Molecular biology and evolution* **34(3)**:772-773. DOI:
485 10.1093/molbev/msw260
486

Laslett D, Canbäck B. 2008. ARWEN, a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* **24**:172-175. DOI: 10.1093/bioinformatics/btm573.

McDonough MM, Ferguson AW, Dowler RC, Gompper ME, Maldonado JE. 2022. Phylogenomic systematics of the spotted skunks (Carnivora, Mephitidae, Spilogale): additional species diversity and Pleistocene climate change as a major driver of diversification. *Molecular Phylogenetics and Evolution* **167**:107266. DOI: 10.1016/j.ympev.2021.107266

Md-Zain BM, Aifat NR, Abdul-Latiff MAB, Mohd-Yusof NS, Ampeng A. 2019. Molecular systematic position of the Sarawak Malay badger, *Mydaus javanensis*. *Malaysian Applied Biology* **48(3)**:125-132.

Minh BQ, Nguyen MAT, Von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Molecular biology and evolution* **30(5)**:1188-1195. DOI: 10.1093/molbev/mst024

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees in Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA. 1-8. DOI: 10.1109/GCE.2010.5676129

Montaña-Lozano P, Moreno-Carmona M, Ochoa-Capera M, Medina NS, Boore JL, Prada CF. 2022. Comparative genomic analysis of vertebrate mitochondrial reveals a differential of rearrangements rate between taxonomic class. *Scientific Reports* **12**:5479. DOI: 10.1038/s41598-022-09512-2.

Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution* **32(1)**:268-274. DOI: 10.1093/molbev/msu300

Ojala D, Montoya J, Attardi G. 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature **290***:470-474. DOI: 10.1038/290470a0.

Parr RL, Maki J, Reguly B, Dakubo GD, Aguirre A, Wittock R, Robinson K, Jakupciak JP, Thayer RE. 2006. The pseudo-mitochondrial genome influences mistakes in heteroplasmy interpretation. *BMC Genomics* **7**:185 (2006). DOI: 10.1186/1471-2164-7-185.

Perna NT, Kocher TD. 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *Journal of Molecular Evolution* **41(3)**:353-358. DOI: 10.1007/BF00186547. PMID: 7563121.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic biology* **67(5)**:901-904. DOI: 10.1093/sysbio/syy032

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology* **61(3)**:539-542. DOI: 10.1093/sysbio/sys029

Samejima H, Meijaard E, Duckworth JW, Yasuma S, Hearn AJ, Ross J, Mohamed A, Alfred R, Bernard H, Boonratana R, Pilgrim JD, Eaton J, Belant JL, Kramer-Schadt S, Semiadi G, Wilting A. 2016. Predicted distribution of the Sunda stink-badger *Mydaus javanensis* (Mammalia: Carnivora: Mephitidae) on Borneo. *Raffles Bulletin of Zoology* **33**:61-70.

Sato JJ, Hosoda T, Wolsan M, Tsuchiya K, Yamamoto M, Suzuki H. 2003. Phylogenetic relationships and divergence times among mustelids (Mammalia: Carnivora) based on nucleotide sequences of the nuclear interphotoreceptor retinoid binding protein and mitochondrial cytochrome b genes. Zoological Science **20(2)**:243-264. DOI: 10.2108/zsj.20.243.

Signes A, Fernandez-Vizarra E. 2018. Assembly of mammalian oxidative phosphorylation complexes I-V and supercomplexes. *Essays in Biochemistry* **62(3)**:255-270. DOI: 10.1042/EBC20170098.

Shokolenko I, Alexeyev M. 2022. Mitochondrial DNA: Consensus and Controversies. *DNA* **2**:131-148. DOI: 10.3390/dna2020010.

Skorupski J. 2022. Characterisation of the complete mitochondrial genome of critically endangered *Mustela lutreola* (Carnivora: Mustelidae) and its phylogenetic and conservation implications. *Genes* **13(1)**:125. DOI: 10.3390/genes13010125.

Sun Y, Zhu Y, Chen C, Zhu Q, Zhu Q, Zhou Y, Zhou X, Zhu P, Li J, Zhang H. 2020. The complete mitochondrial genome of *Dysgonia stuposa* (Lepidoptera: Erebidae) and phylogenetic relationships within Noctuoidea. *PeerJ* **8(3)**:e8780. DOI: 10.7717/peerj.8780

Watanabe YI, Suematsu T, Ohtsuki T. 2014. Losing the stem-loop structure from metazoan mitochondrial tRNAs and co-evolution of interacting factors. *Frontiers in Genetics* **5**:109. DOI: 10.3389/fgene.2014.00109

Wilkinson GS, Mayer F, Kerth G, Petri B. 1997. Evolution of repeated sequence arrays in the D-loop region of bat mitochondrial DNA. *Genetics* **146(3)**:1035-1048.

Wong ST, Belant JL, Sollmann R, Mohamed A, Niedballa J, Mathai J, Meijaard E, Street GM, Kissing J, Mannan S, Wilting A. 2017. Habitat associations of the Sunda stink-badger *Mydaus javanensis* in three forest reserves in Sabah, Malaysian Borneo. *Mammalian Biology* **88**:75-80. DOI: 10.1016/j.mambio.2017.11.010.

Table 1(on next page)

Details of the PCR primers and conditions used in the amplification of the Sunda stink-badger (*Mydaus javanensis*) mitogenome.

Table 1 Details of the PCR primers and conditions used in the amplification of the Sunda stink-badger (*Mydaus javanensis*) mitogenome.

Primer name		Sequence (5'-3')	Amplification size	Annealing temperature used (°C)
Leo16SLRpcr_F	Forward	CAGGACATCCCGATGGTGCAG	~16 kb	60
Leo16SLRpcr_R	Reverse	ATCCAACATCGAGGTCGTAAAC		
MJ19F	Forward	TGAAATTGACCTCCCCGTGA	688 bp	55
MJ19R	Reverse	AGGCGCCTTTAGACTAACAGA		

Table 2(on next page)

Availability of mitogenome and *cytb* gene sequences for the 12 extant species in the family Mephitidae.

No mitogenome sequences were available for *C. humboldtii*, *M. macroura*, *M. javanensis* and *M. marchei*. Of these four, *cytb* gene sequences were only available for *M. macroura* and *M. javanensis*.

Table 2 Availability of mitogenome and *cyt b* gene sequences for the 12 extant species in the family Mephitidae. No mitogenome sequences were available for *C. humboldtii*, *M. macroura*, *M. javanensis* and *M. marchei*. Of these four, *cytb* gene sequences were only available for *M. macroura* and *M. javanensis*.

Genera	Species	Availability of Mitogenome sequences in GenBank	Availability of *cytb* gene sequences in GenBank
Conepatus	*C. chinga*	NC_042596 (complete)	Available
	C. humboldtii	None	None
	C. leuconotus	MW205848 (partial genome)	Available
	C. semistriatus	MW205849 (partial genome)	Available
Mephitis	*M. macroura*	None	KY026063
	M. mephitis	NC_020648 (complete)	Available
Mydaus	*M. javanensis*	None	AB564095
	M. marchei	None	None
Spilogale	*S. angustifrons*	MW205870 (complete)	Available
		MW205885 *S. angustifrons yucatanensis* (complete)	Available
	S. gracilis	MW205896 *S. gracilis gracilis* (complete)	Available
		MW205880 *S. gracilis leucoparia* (complete)	Available
		MW205868 *S. gracilis martirensis* (complete)	Available
		MW205862 *S. gracilis lucasana* (complete)	Available
	S. putorius	NC_010497 (complete)	Available
		MW205890 *S. interrupta* (complete)	Available
		S. putorius putorius (None)	MG753651
		S. putorius ambarvalis (None)	MG753655
	S. pygmaea	MW205863 (partial)	Available

Table 3(on next page)

Composition and annotation of the newly sequenced mitogenome of the Sunda stink-badger (*Mydaus javanensis)*.

Table 3 Composition and annotation of the newly sequenced mitogenome of the Sunda stink-badger (*Mydaus javanensis)*.

	Feature name	Type	Start position	End position	Size	Direction	Start codon	Stop codon	Intergenic nucleotides
1	*trnF*	tRNA	1	69	69	Forward	-	-	0
2	*rrnS*	rRNA	70	1,026	957	Forward	-	-	0
3	*trnV*	tRNA	1,027	1,094	68	Forward	-	-	0
4	*rrnL*	rRNA	1,095	2,666	1,572	Forward	-	-	0
5	*trnL2*	tRNA	2,667	2,741	75	Forward	-	-	3
6	*ND1*	gene	2,745	3,700	956	Forward	ATG	TA(A)	0
7	*trnI*	tRNA	3,701	3,769	69	Forward	-	-	-3
8	*trnQ*	tRNA	3,767	3,839	73	Reverse	-	-	1
9	*trnM*	tRNA	3,841	3,910	70	Forward			0
10	*ND2*	gene	3,911	4,952	1,042	Forward	ATA	T(AA)	0
11	*trnW*	tRNA	4,953	5,019	67	Forward	-	-	13
12	*trnA*	tRNA	5,033	5,100	68	Reverse	-	-	1
13	*trnN*	tRNA	5,102	5,174	73	Reverse	-	-	0
14	O_L	rep_origin	5,175	5,207	33	Forward	-	-	-1
15	*trnC*	tRNA	5,207	5,272	66	Reverse	-	-	0
16	*trnY*	tRNA	5,273	5,339	67	Reverse	-	-	1
17	*COI*	gene	5,341	6,885	1,545	Forward	ATG	TAA	-3
18	*trnS2*	tRNA	6,883	6,951	69	Reverse	-	-	5
19	*trnD*	tRNA	6,957	7,023	67	Forward	-	-	0
20	*COII*	gene	7,024	7,707	684	Forward	ATG	TAA	3
21	*trnK*	tRNA	7,711	7,779	69	Forward	-	-	1
22	*ATP8*	gene	7,781	7,984	204	Forward	ATG	TAA	-43
23	*ATP6*	gene	7,942	8,622	681	Forward	ATG	TAA	-1
24	*COIII*	gene	8,622	9,405	784	Forward	ATG	T(AA)	0
25	*trnG*	tRNA	9,406	9,474	69	Forward	-	-	0
26	*ND3*	gene	9,475	9,821	347	Forward	ATA	TA(A)	0
27	*trnR*	tRNA	9,822	9,889	68	Forward	-	-	0
28	*ND4L*	gene	9,890	10,186	297	Forward	ATG	TAA	-7
29	*ND4*	gene	10,180	11,557	1,378	Forward	ATG	T(AA)	0
30	*trnH*	tRNA	11,558	11,626	69	Forward	-	-	0
31	*trnS1*	tRNA	11,627	11,685	59	Forward	-	-	0
32	*trnL1*	tRNA	11,686	11,755	70	Forward	-	-	0
33	*ND5*	gene	11,756	13,576	1,830	Forward	ATT	TAA	-17
34	*ND6*	gene	13,560	14,087	528	Reverse	ATG	TAA	0
35	*trnE*	tRNA	14,088	14,156	69	Reverse	-	-	4
36	*CYTB*	gene	14,161	15,300	1,140	Forward	ATG	AGA	0
37	*trnT*	tRNA	15,301	15,369	69	Forward	-	-	-1
38	*trnP*	tRNA	15,369	15,434	66	Reverse	-	-	-40
39	Control region	D-loop	15,395	16,391	997	Forward	-	-	0

Table 4(on next page)

Nucleotide composition and skewness of the mitogenome in the Sunda stink-badger (*Mydaus javenensis*).

Table 4 Nucleotide composition and skewness of the mitogenome in the Sunda stink-badger (*Mydaus javenensis*).

Regions	Size	A%	T%	C%	G%	A+T %	C+G %	AT skew	GC skew
Whole genome	16,391	33.9	28.0	24.9	13.2	61.9	38.1	0.095	-0.307
PCGs	11,416	32.4	29.6	25.4	12.6	62.0	38.0	0.045	-0.337
tRNA genes	1,509	33.3	31.3	16.4	19.0	64.6	35.4	0.031	0.073
rRNA genes	2,529	37.1	24.7	21.0	17.2	61.8	38.2	0.201	-0.099
Control region	997	30.4	28.1	26.0	15.5	58.5	41.5	0.039	-0.253

Table 5(on next page)

Details of tandem repeats within control regions of mephitids

Details of tandem repeat regions detected within mitochondrial control regions of mephitids

Table 5 Details of tandem repeat regions detected within mitochondrial control regions of mephitids

Species (Genbank accession number)	Gene size (bp)	No. of tandem repeat regions	Tandem repeat positions in control region
Mydaus javanensis MJ19 (OP442081)	997	1	a. Right domain (Between CSB-1 & CSB-2)
Spilogale gracilis leucoparia (MW205880)	853	None	None
Spilogale gracilis gracilis (MW205896)	819	None	None
Spilogale putorius (NC_010497)	1138	2	a. Right domain (Between CSB-1 & CSB-2) b. Left Domain
Mephitis mephitis (NC_020648)	1103	1	a. Right domain (Between CSB-1 & CSB-2)
Spilogale augustifrons (MW205870)	899	1	a. Left domain
*Conepatus chinga** (NC_042596)	1067	1	a. Right domain (Between CSB-1 & CSB-2)
*Conepatus leuconotus leuconotus** (AY159816)	1218	1	a. Right domain (Between CSB-1 & CSB-2)

Note: The asterisk (*) denotes that the sequence is treated as a reference, according to Ketmaier & Bernardini (2005).

Figure 1

Circular map of the Sunda stink-badger (*Mydaus javanensis*) mitochondrial genome.

The protein coding and ribosomal genes are shown in standard abbreviations. Different colors represent the different gene blocks or regions in the mitogenome.

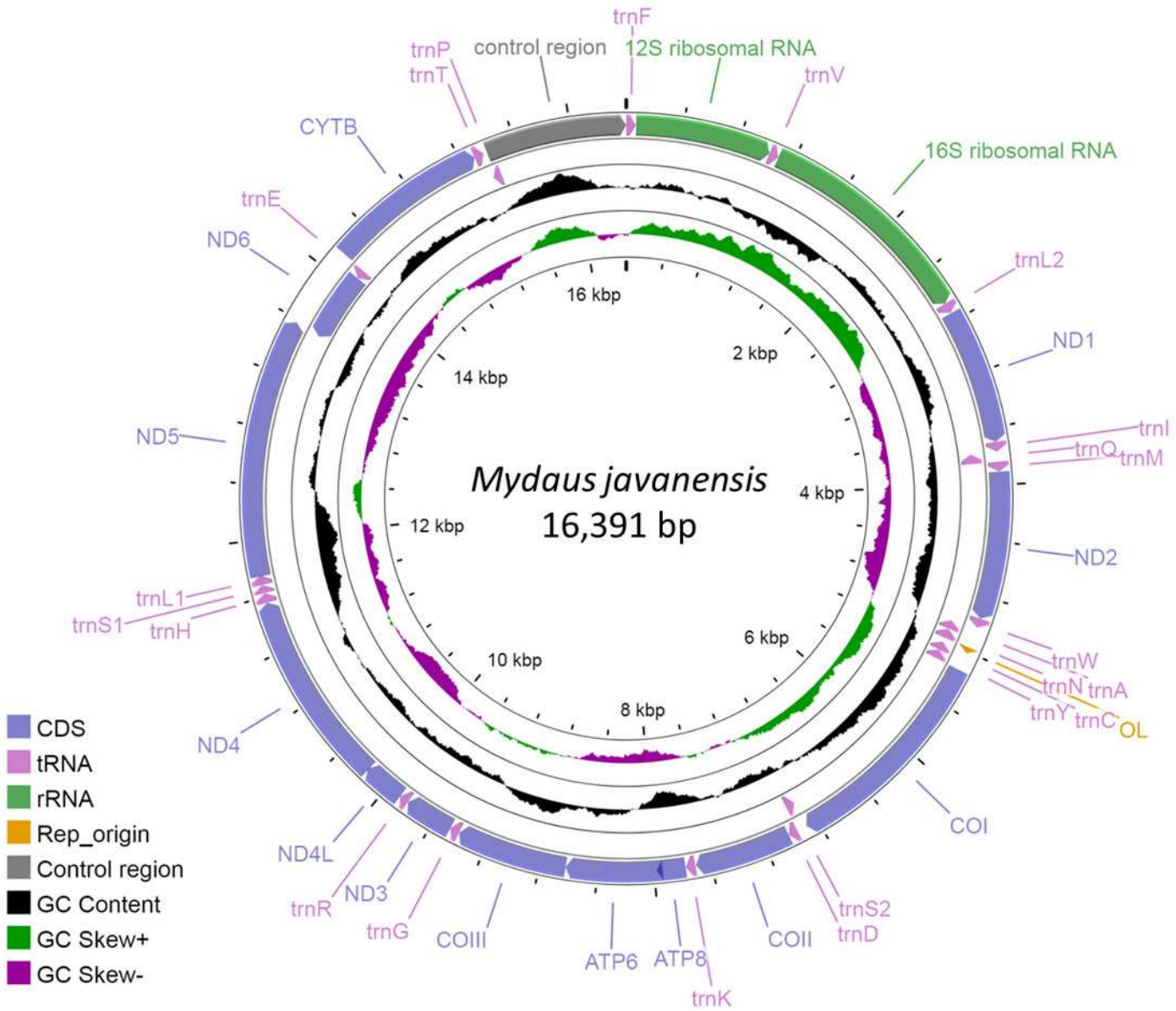

Figure 2

The codon number and relative synonymous codon usage

The codon number and relative synonymous codon usage (RSCU) of the 13 protein coding genes (PCGs) in *Mydaus javanensis* mitogenome.

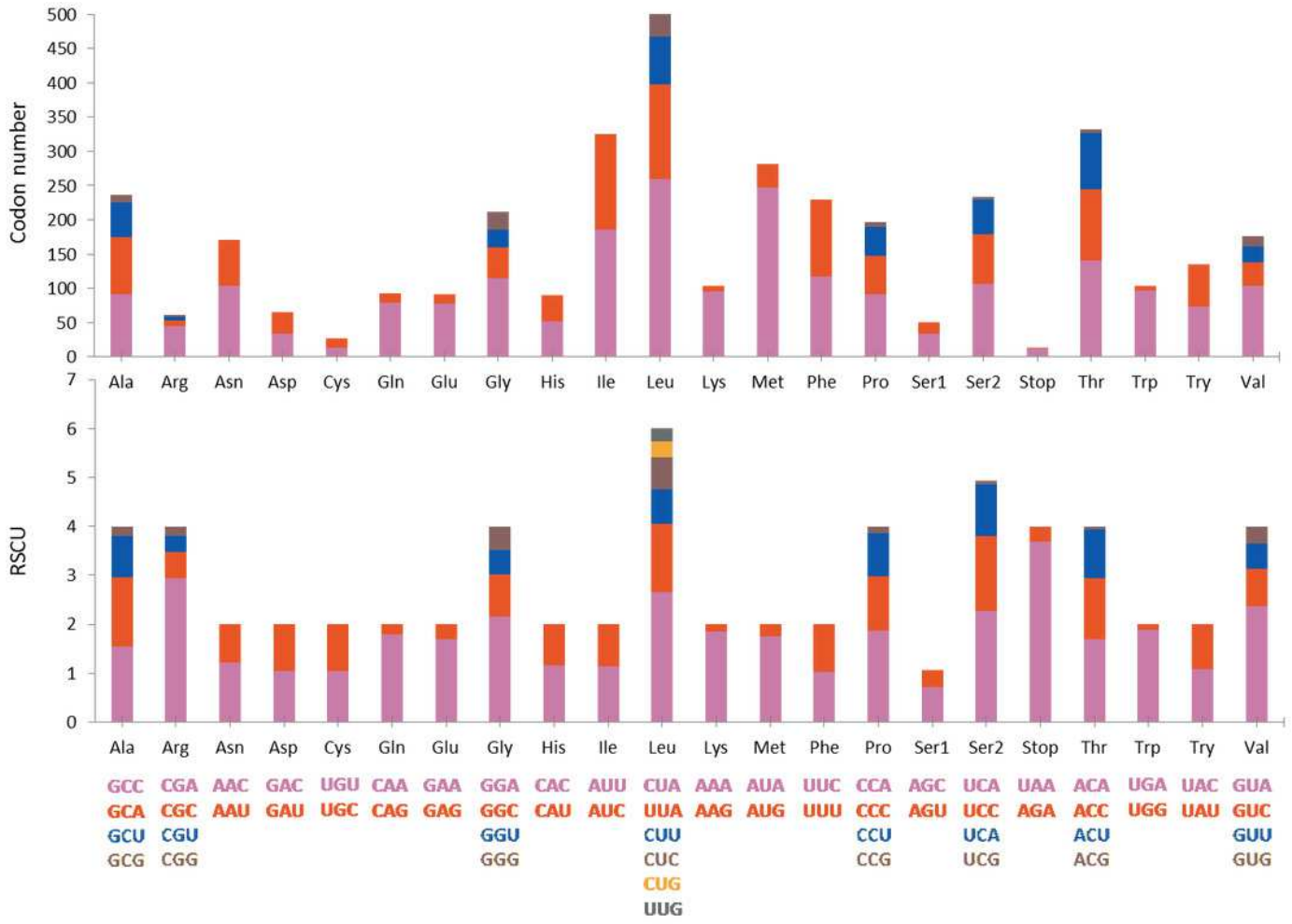

Figure 3

Putative secondary structures of Mydaus javanensis tRNAs indicating the typical clover leaf shape

Putative secondary structures of *Mydaus javanensis* tRNAs indicating the typical clover leaf shape consisting of the acceptor stem, DHU loop, Anticodon loop, TΨC Loop, variable arm and the discriminator base. The tRNAs are labeled with their corresponding amino acids. The dashes (-) indicate the Watson-Crick bonds and the dot (⋆) indicate a mispairing between the nucleotides.

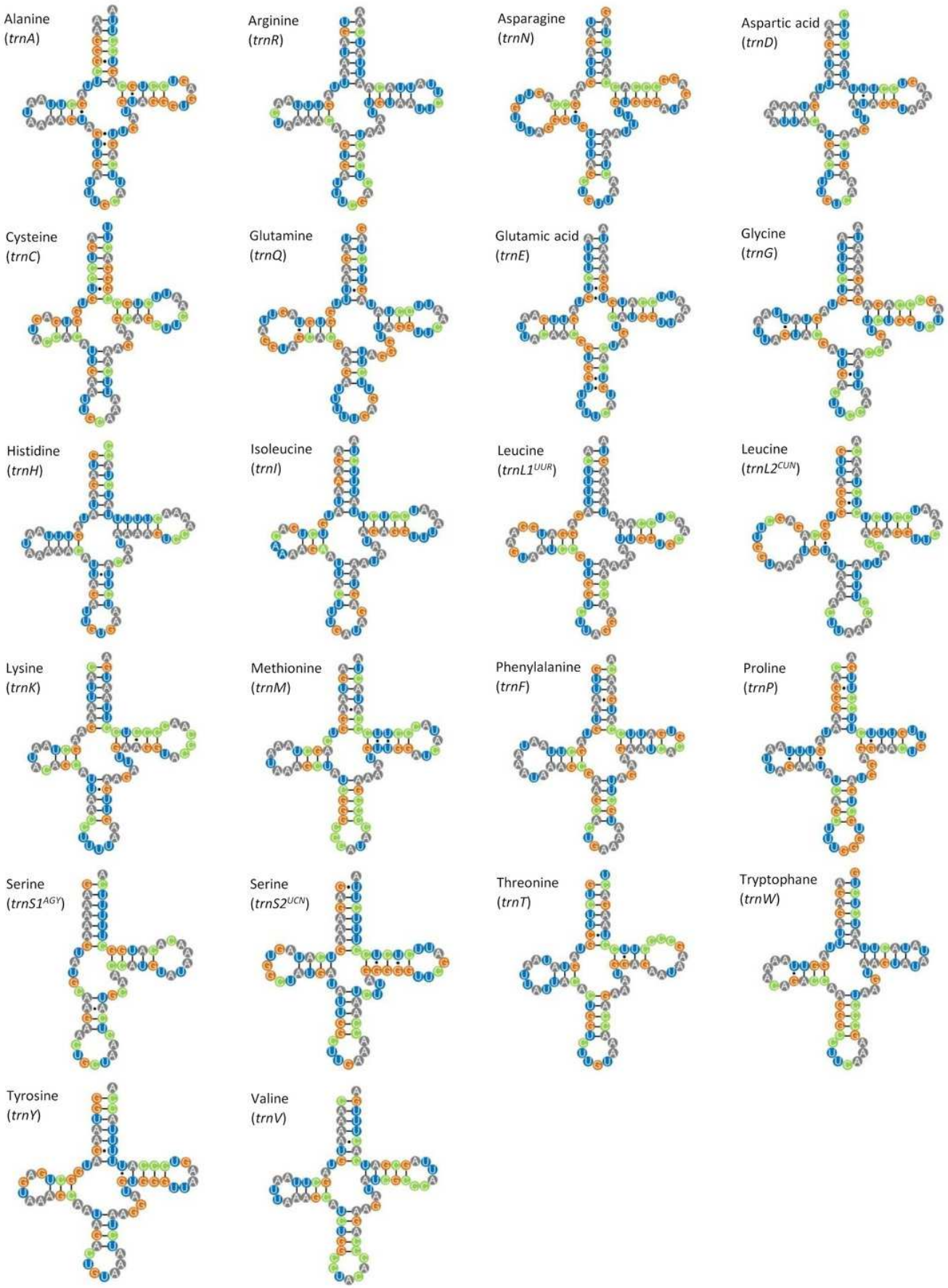
Alanine (trnA)
Arginine (trnR)
Asparagine (trnN)
Aspartic acid (trnD)
Cysteine (trnC)
Glutamine (trnQ)
Glutamic acid (trnE)
Glycine (trnG)
Histidine (trnH)
Isoleucine (trnI)
Leucine (trnL1UUR)
Leucine (trnL2CUN)
Lysine (trnK)
Methionine (trnM)
Phenylalanine (trnF)
Proline (trnP)
Serine (trnS1AGY)
Serine (trnS2UCN)
Threonine (trnT)
Tryptophane (trnW)
Tyrosine (trnY)
Valine (trnV)

Figure 4

Schematic diagram of the organization of the mitochondrial control region (CR) of *Mydaus javanensis*

Schematic diagram of the organization of the mitochondrial control region (CR) of *Mydaus javanensis*. The CR is shown to bound by the $tRNA^{Pro}$ and $tRNA^{Phe}$ genes. The termination of the displacement loop (D-loop) motif (D-loop termination; blue block), termination-associated sequence (TAS-A; orange block) and conserved sequence blocks (CSBs; green blocks) were specified according to Ketmaier and Bernardini (2005). The grey striped region indicates that the repeat region, RS3 (position: 607-727) is located between CSB1 and CSB2.

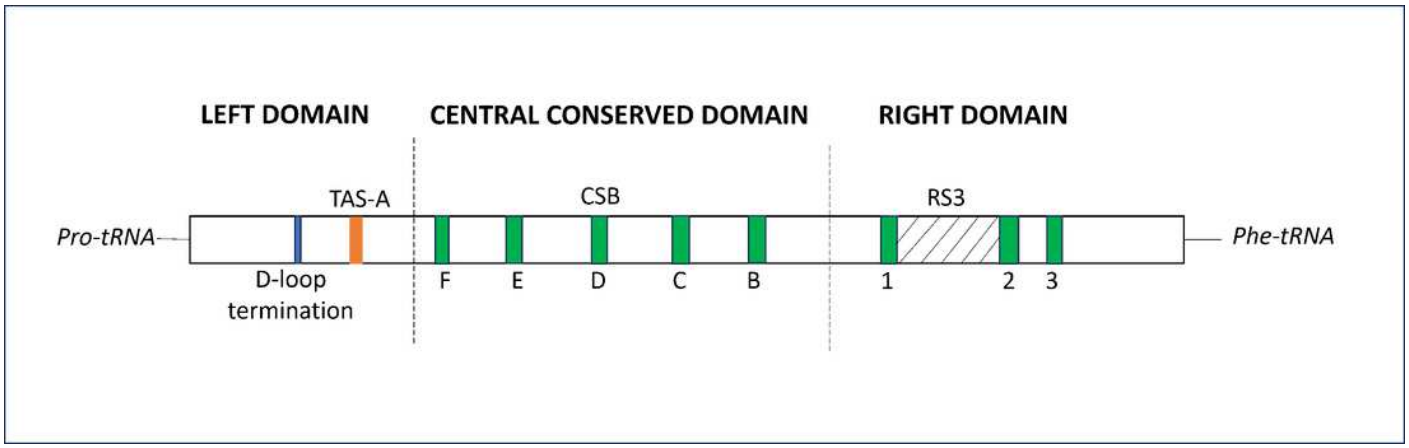

Figure 5

Alignment of the D-loop termination, TAS-A, and eight CSBs of mitochondrial control region (CR) of *Mydaus javanensis* and other related species

Alignment of the D-loop termination, TAS-A, and eight CSBs of mitochondrial control region (CR) of *Mydaus javanensis* and other related species included in the study. The sequences of *Conepatus chinga*, *Conepatus leuconotus leuconotus* and *Lutra lutra* serve as references for the alignments, according to Ketmaier and Bernardini (2005). The black background indicates conserved bases while white background indicates variations observed in the nucleotides. Numbers above the alignment point out the first nucleotide positions of each region based on mitochondrial control region of *Mydaus javanensis*.

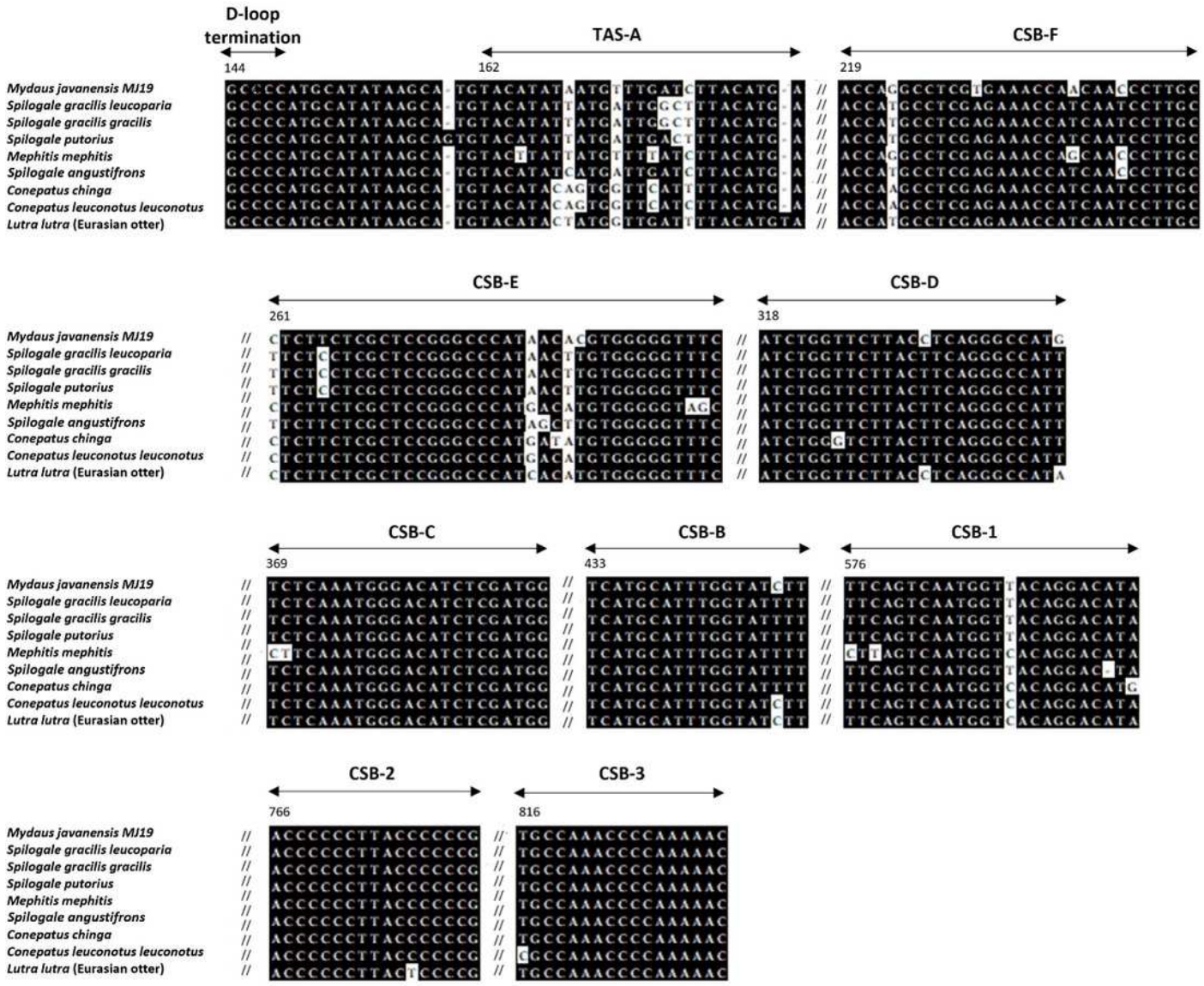
D-loop termination
TAS-A
CSB-F
144
162
219
Mydaus javanensis MJ19
Spilogale gracilis leucoparia
Spilogale gracilis gracilis
Spilogale putorius
Mephitis mephitis
Spilogale angustifrons
Conepatus chinga
Conepatus leuconotus leuconotus
Lutra lutra (Eurasian otter)
CSB-E
CSB-D
261
318
CSB-C
CSB-B
CSB-1
369
433
576
CSB-2
CSB-3
766
816

Figure 6

Maximum Likelihood (ML) phylogenetic tree of the Sunda stink-badger

Maximum Likelihood (ML) phylogenetic tree of the Sunda stink-badger (*Mydaus javanensis*) in comparison with mitogenomes of Mephitidae and other selected members of suborder Caniformia. The tree was based on the concatenated sequences of 13 Protein Coding Genes (PCGs). The clouded leopard (*Neofelis nebulosa*) was used as an outgroup. The sequence characterized in this study is highlighted in red. Genera of the mephitids are indicated in blue highlights. The bootstrap values of the branches were displayed at each node.

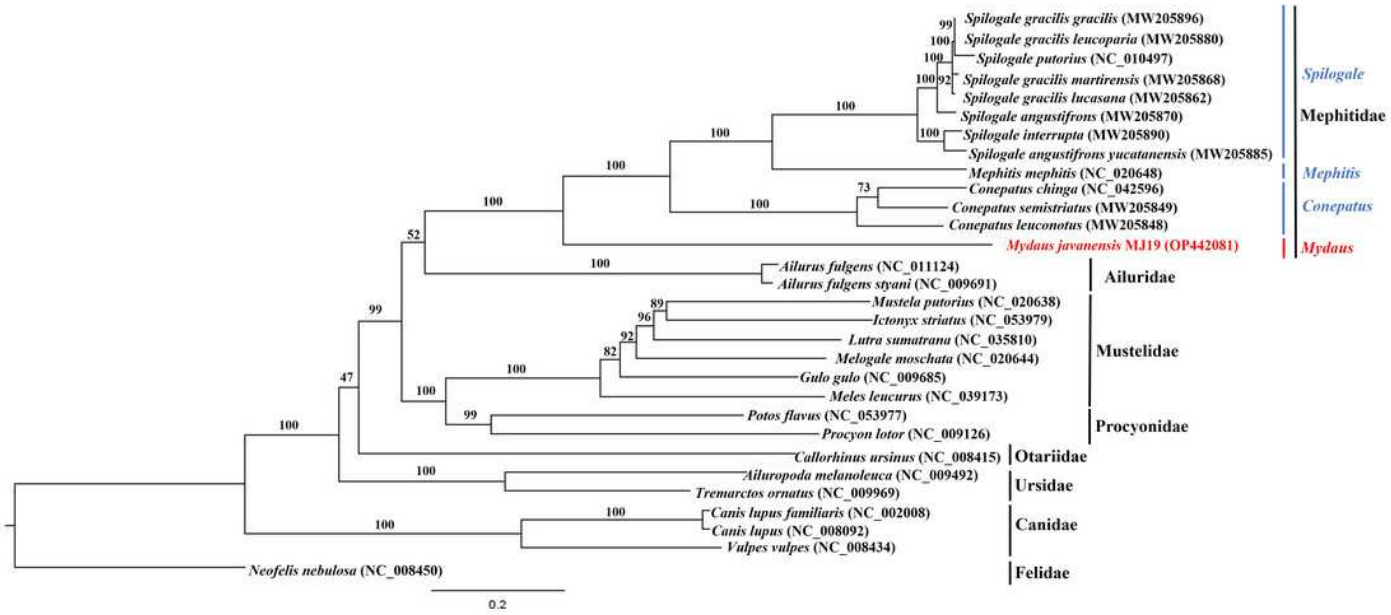

Figure 7

Bayesian inference (BI) phylogenetic tree of the Sunda stink-badger

Bayesian Inference (BI) phylogenetic tree of the Sunda stink-badger (*Mydaus javanensis*) in comparison with mitogenomes of Mephitidae and other selected members of suborder Caniformia. The tree was based on the concatenated sequences of 13 Protein Coding Genes (PCGs). The clouded leopard (*Neofelis nebulosa*) was used as an outgroup. The sequence characterized in this study is highlighted in red. Genera of the mephitids are indicated in blue highlights. The probability values of the branches were displayed at each node.

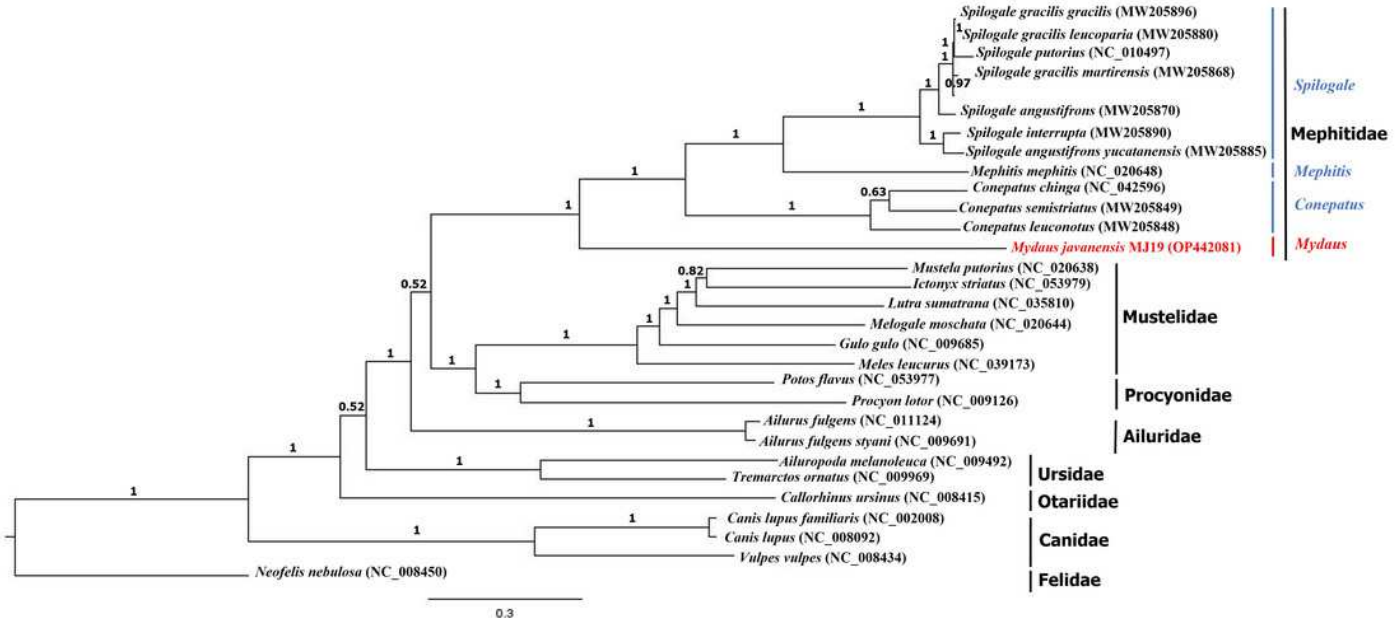

Figure 8

Maximum Likelihood (ML) Phylogenetic tree based on *cyt b* gene of the Sunda stink-badger

Maximum Likelihood (ML) phylogenetic tree of the Sunda stink-badger (*Mydaus javanensis*) in comparison with mitogenomes of Mephitidae and other selected members of suborder Caniformia. The tree was based on the *cyt b* gene. The clouded leopard (*Neofelis nebulosa*) was used as an outgroup. The sequence characterized in this study is highlighted in red. Genera of the mephitids are indicated in blue highlights. The bootstrap values of the branches were displayed at each node.

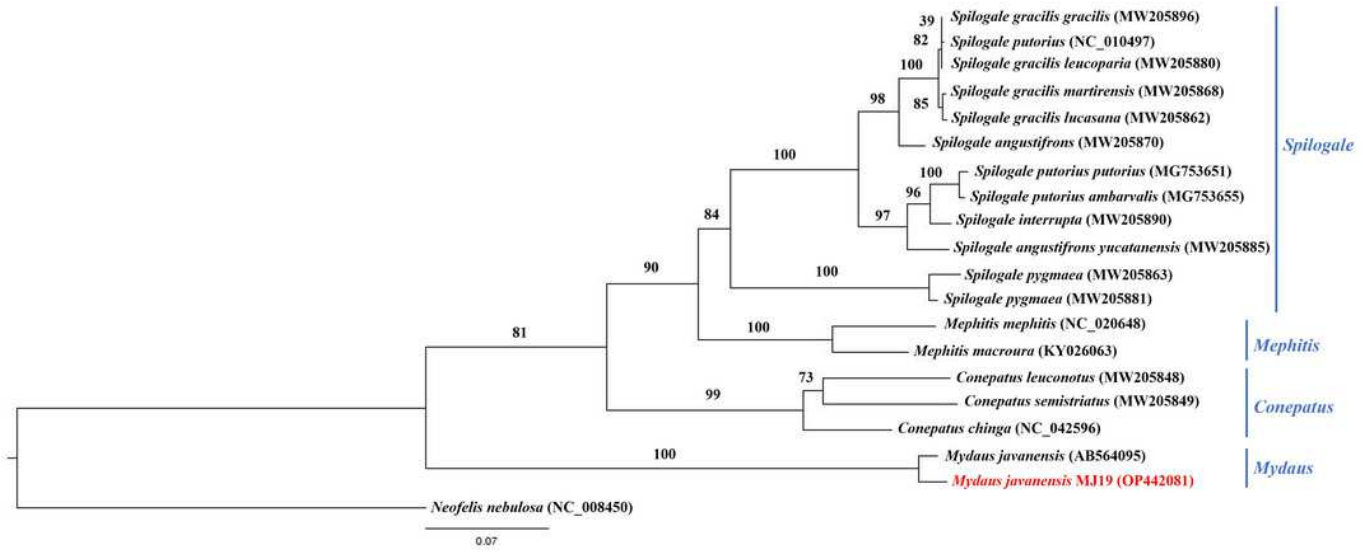

Figure 9

Bayesian inference (BI) phylogenetic tree based on the *cyt b* gene of the Sunda stink-badger

Bayesian Inference (BI) phylogenetic tree of the Sunda stink-badger (*Mydaus javanensis*) in comparison with mitogenomes of Mephitidae and other selected members of suborder Caniformia. The tree was based on the *cyt b* gene. The clouded leopard (*Neofelis nebulosa*) was used as an outgroup. The sequence characterized in this study is highlighted in red. Genera of the mephitids are indicated in blue highlights. The probability values of the branches were displayed at each node.