

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 **Title: An expanded mammal mitogenome dataset from Southeast Asia**

2 Faezah Mohd Salleh^{1,2*}, Jazmín Ramos-Madrigal¹, Fernando Peñaloza^{3,4}, Shanlin Liu^{1,5}, Mikkel-Holger S.
3 Sinding^{1,6}, Riddhi P. Patel^{3,7}, Renata Martins³, Dorina Lenz³, Christian Roos⁸, Mohd Shahir Shamsir²,
4 Mohammad Shahfiz Azman⁹, Burton K. Lim¹⁰, Stephen J. Rossiter¹¹, Andreas Wilting³, M. Thomas P.
5 Gilbert^{1,12*}

6 *Correspondence: faezah@fbb.utm.my and tgilbert@snm.ku.dk

7
8 ¹Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster
9 Voldgade 5-7, 1350, Copenhagen, Denmark

10 ²Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru,
11 Johor, Malaysia

12 ³Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke Strasse 17, 10315 Berlin, Germany.

13 ⁴Undergraduate Program on Genomic Sciences, Universidad Nacional Autonoma de Mexico, 62210
14 Cuernavaca, Mexico

15 ⁵BGI-Shenzhen, Shenzhen, GuangDong, China

16 ⁶Natural History Museum, University of Oslo, P.O. Box 1172 Blindern, NO-0318, Oslo, Norway

17 ⁷Freie Universität Berlin, Kaiserswerther Str. 16-18, 14195 Berlin, Germany

18 ⁸Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center, Leibniz Institute for
19 Primate Research, Kellnerweg 4, 37077, Göttingen, Germany

20 ⁹Forest Biodiversity Division, Forest Research Institute Malaysia, 52109 Kepong, Selangor, Malaysia.

21 ¹⁰Department of Natural History, Royal Ontario Museum, Toronto, Canada

22 ¹¹School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London
23 E1 4NS, United Kingdom

24 ¹²NTNU University Museum, Norwegian University of Science and Technology, Trondheim, Norway

25

26

27

28

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

29 Abstract

30 Background:

31 Southeast (SE) Asia is one of the most biodiverse regions in the world and it holds approximately 20% of
32 all mammal species. Despite this, the majority of South East (SE) Asia's genetic diversity is still poorly
33 characterised. The growing interest in using environmental DNA (eDNA) to assess and monitor SE Asian
34 species, in particular threatened mammals - has created the urgent need to expand the available
35 reference database of mitochondrion barcode and complete mitogenome sequences. We have
36 contributed to this need, by generating 75 new mitogenome sequences reconstructed from DNA isolated
37 from a range of historical and modern tissue samples.

38 Findings:

39 Approximately 56 gigabases of raw sequence were generated. From this data we assembled 75 complete
40 mitogenome sequences, with an average depth of coverage of 102.9x and 55.2x for modern samples and
41 historical samples, respectively. This dataset represents 55 species, of which 32 species have currently
42 no mitogenome available. The mitogenomes were geotagged to their sampling location, where known, to
43 display a clearer geographical distribution of the species.

44 Conclusion:

45 Our new database of 55 taxa will greatly enhance the utility of eDNA approaches for monitoring mammals
46 in SE Asia, as it greatly increases the likelihoods that identification of metabarcoding sequencing reads
47 can be assigned to reference sequences. This magnifies the confidence in species detections and thus
48 allows more robust surveys and monitoring programmes of SE Asia's threatened mammal biodiversity.
49 The extensive collections of historical samples from SE Asia in western and SE Asian museums should
50 serve as additional valuable material to further enrich this reference database.

51 Keywords:

52 invertebrate-derived (iDNA), metabarcoding, Malaysia, GenBank, Taxonomic assignment

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

57 Data description:

58 Context

59 Southeast (SE) Asia is one of the most biodiverse regions in the world, hosting ~20% of mammal species,
60 but is experiencing rapid deforestation for agriculture and development. To assess the ecological
61 consequences of land-use change, there is growing interest in using environmental DNA (eDNA) to
62 monitor mammal populations, particularly threatened taxa that often underpin conservation policies. Yet
63 current efforts are hampered by the lack of a reference database of mitochondrion barcodes and
64 complete mitogenome sequences. We have contributed to this need, by generating 75 new mitogenome
65 sequences reconstructed from DNA isolated from a range of historical and modern tissue samples.

67 DNA extraction

68 Genomic DNA was extracted from different sample types of 75 small mammals, comprising 55 species,
69 listed in [Table 1](#) and [Table 2](#). DNA from modern tissue and blood samples was isolated using the Qiagen
70 DNeasy extraction kit (Qiagen, Hilden, Germany) or Invitex DNA extraction kit (Invitex GmbH, Berlin,
71 Germany), as per standard protocols. Historical samples obtained from the Zoological Museum, Natural
72 History Museum of Denmark, University of Copenhagen (ZM, KU) were treated differently according to
73 type of tissue ([Additional file 1a](#)), while at the German Primate Center DNA extraction from museum
74 specimens followed Liedigk et al. (2015) [1]. Complete details of sample information are provided in
75 [Additional file 2](#).

77 Mitogenome sequencing, assembly and annotation

78 Mitogenomes were generated using several approaches. In Copenhagen, author FMS constructed
79 Illumina shotgun libraries with insert sizes ranging between 50 bp to 400 bp. To construct libraries, DNA
80 was sheared to the target size range using Bioruptor® XL (Diagenode, USA) and converted into an
81 Illumina-compatible sequencing library using the NEBNext E6070 kit (New England Biolabs, UK). The
82 libraries were PCR amplified with index primers and purified using Qiaquick columns (Qiagen, Hilden,
83 Germany) according to the manufacturer's instruction ([Additional file 1b](#)). Multiple libraries were combined
84 together into three pools, normalized to 10 nM and sequenced across three lanes of Illumina HiSeq 2500

1
2
3
4 85 using SR100bp chemistry. In Berlin and Goettingen, mitogenomes were generated by author PRP and
5
6 86 CR using overlapping PCR products using long range PCR ([Additional file 1c](#)) followed by library
7
8 87 construction and MiSeq sequencing, or Sanger sequencing as described in [2] and [1, 3, 4], respectively.
9
10 88 Author RM's mitogenomes were done using methods outlined in [5].
11
12 89
13
14 90 Raw reads for FMS samples were assembled independently by authors FMS and FP using two different
15
16 91 approaches, then compared for consistency. Author FMS trimmed the reads for sequencing adapters, low
17
18 92 quality stretches and leading/tailing Ns using AdapterRemoval 1.2 [6]. The mitochondrial genome was
19
20 93 reconstructed with MITObim v1.8 [7] using the reference mitogenome of the closest species available in
21
22 94 GenBank as the seed reference ([Additional file 2](#)). In order to obtain the mapping statistics of the
23
24 95 samples, we ran PALEOMIX v 1.2.6 [8] with default parameters where reads shorter than 25 bp after
25
26 96 trimming were discarded. The trimmed reads were aligned against the newly assembled mitogenome
27
28 97 generated by MITObim using Burrows–Wheeler Aligner (BWA) [9]. Alignments showing low-quality scores
29
30 98 and PCR duplicates were further removed using the MarkDuplicates program from Picard tools, and
31
32 99 reads were locally realigned around small insertions and deletions (indels) to improve overall genome
33
34 100 quality using the IndelRealigner tool from the Genome Analysis Toolkit (GATK)[10]. In contrast, author
35
36 101 FP inputted the trimmed reads into mitoMaker (<https://sourceforge.net/projects/mitomaker/>), which
37
38 102 performs a de-novo and reference-based assembly using SOAPdenovoTrans v1.03 [11] and MITObim
39
40 103 v1.7 [7]. Post assembly, the FMS and FP mitogenomes were manually compared for consistency by FMS
41
42 104 to generate the final consensus sequences. These assemblies were automatically annotated using
43
44 105 tRNAscan-SE v1.4 [12] and BLAST v2.2.29 [13] using the mitochondrial genomes found in the NCBI
45
46 106 RefSeq (<https://www.ncbi.nlm.nih.gov/refseq/>) as references.
47
48
49 107
50
51 108 For mitogenome constructed by author RM, Illumina sequence reads were de-multiplexed according to
52
53 109 the respective indexes with the Illumina software bcl2fastq v2.17 (Illumina, San Diego, CA, USA) and
54
55 110 adapters were clipped from the sequence reads with the software cutadapt v1.3 [14]. Quality trimming
56
57 111 was done through a sliding window approach (10 bp; Q20) and all reads shorter than 20 bp were
58
59 112 removed from the analyses. Mitogenome references from target or closely related species were used for
60
61
62
63
64
65

1
2
3
4 113 mapping of the sequencing reads. Aligned reads were de-duplicated using MarkDuplicates from Picard-
5
6 114 tools v1.106 (<https://github.com/broadinstitute/picard>). VariantCalling was carried out using Samtools v1.1
7
8
9 115 [9] and Bcftools v1.2 (<http://github.com/samtools/bcftools>). For each sample, GATK [10] Variant Calling
10
11
12 116 output files were further filtered to have a minimum read coverage $\geq 3x$, and variants were only called
13
14
15 117 when the corresponding base was represented by $\geq 50\%$, otherwise this position was "N"-masked.

16
17 118
18
19
20 119 Numbers of raw reads generated for each samples and mapping statistics for all 75 mitogenome
21
22 120 assemblies are shown in [Additional file 2](#). Sanger sequenced mitogenomes were checked with 4Peaks
23
24 121 1.8 (www.nucleobytes.com), assembled with SeaView 4.5.4 [15] and annotated with DOGMA [16]. All
25
26 122 mitogenomes were checked manually by eye to identify possible errors caused by insertion and deletions.
27
28 123 The final mitochondrial genomes have been uploaded to GenBank (accession numbers are provided in
29
30 124 Table 1 and Table 2). The details of all new mitogenomes assembled in this work are given in Table 1
31
32 125 and Table 2. Mitogenomes (62 samples) with known localities were geotagged and mapped to display its
33
34 126 geographical distribution (Figure 1).

35
36 127
37
38 128 **Phylogenetic analysis**
39
40 129 All the sequenced mitogenomes were aligned using MAFFT v7.158b [17], using the E-INS-i option
41
42 130 ([Additional file 3](#)). RAxML v8.0.26 [18] was used to perform the phylogenetic analysis with a
43
44 131 GTR+GAMMA model of nucleotide substitution. To obtain node support, we used 100 bootstrap pseudo-
45
46 132 replicates (Figure 2). The newick file is provided as [Additional file 4](#).

47
48 133
49
50 134 **Availability of supporting data**
51
52 135 Raw shotgun data are deposited in the SRA under bioproject number PRJNA361218 and is available in the
53
54 136 GigaScience repository, GigaDB <ftp://user27:SallehFaezah@climb.genomics.cn>.

55
56
57 137
58
59 138

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

139

140 Abbreviations

141 SE: South East; eDNA: environmental deoxyribonucleic acid; bp: basepair; PCR: polymerase chain
142 reaction; BWA: Burrows–Wheeler Aligner; GATK: Genome Analysis Toolkit; BLAST: Basic Local
143 Alignment Search Tool; NCBI RefSeq: National Center for Biotechnology Information Reference
144 Sequence Database; MAFFT: Multiple Alignment using Fast Fourier Transform ; RAxML: Randomized
145 Axelerated Maximum Likelihood.

146

147

148 Acknowledgements

149 This project was funded by the Malaysian Government, Lundbeck Foundation (R52-5062), Leibniz-
150 Association grant SAW-2013-IZW-2, the German Federal Ministry of Education and Research grant
151 BMBF FKZ: 01LN1301A and the German Primate Center. We thank the Danish National High-throughput
152 Sequencing Centre for assistance in generating Illumina data. We thank the staff of the zoos in Hannover,
153 Copenhagen, Gettorf, Belfast, Bristol, Howletts Wild Animal Park, the staff of the Ludwig-Maximilians-
154 University Munich, the Bavarian State Collection Munich, the National Museums Scotland, Edinburgh, the
155 Natural History Museum Berlin, Naturalis Leiden, and the Zoological Museum, Natural History Museum of
156 Denmark, University of Copenhagen for providing samples for this work.

157

158 Authors' contributions

159 FMS, AW, and MTPG conceived the project. FMS, MHSS, MSS, MSA, RM, PRP, CR, BKL, and SJR
160 collected the samples and extracted the genomic DNA. FMS, RM, PRP and CR constructed the libraries
161 and did sequencing. FMS, JRM, FP, SL, PRP, RM, DL and CR assembled the mitogenomes and
162 performed mitogenome analysis. FMS, SL, PRP and MTPG wrote the article. All authors discussed the
163 project and data. All authors read and approved the final manuscript.

164

165 Competing interests

166 The authors declare that they have no competing interests.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193

Author details

¹Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350, Copenhagen, Denmark. ²Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia. ³Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke Strasse 17, 10315 Berlin, Germany. ⁴Undergraduate Program on Genomic Sciences, Universidad Nacional Autonoma de Mexico, 62210 Cuernavaca, Mexico. ⁵BGI-Shenzhen, Shenzhen, Guangdong, China. ⁶Natural History Museum, University of Oslo, P.O. Box 1172 Blindern, NO-0318, Oslo, Norway. ⁷Freie Universität Berlin, Kaiserswerther Str. 16-18, 14195 Berlin, Germany. ⁸Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077, Göttingen, Germany. ⁹Forest Biodiversity Division, Forest Research Institute Malaysia, 52109 Kepong, Selangor, Malaysia. ¹⁰Department of Natural History, Royal Ontario Museum, Toronto, Canada. ¹¹School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK. ¹²NTNU University Museum, Norwegian University of Science and Technology, Trondheim, Norway.

1
2
3
4 194
5
6 195
7
8 196
9
10
11 197
12
13 198
14
15
16 199
17 200
18
19 201
20 202
21 203
22
23 204
24 205
25
26 206
27 207
28
29 208
30 209
31
32 210
33 211
34
35 212
36 213
37 214
38
39 215
40 216
41 217
42
43 218
44 219
45
46 220
47 221
48
49 222
50 223
51
52 224
53 225
54
55 226
56 227
57
58 228
59 229
60
61 230
62
63
64
65

References

1. Liedigk R, Kolleck J, Böker KO, et al (2015) Mitogenomic phylogeny of the common long-tailed macaque (*Macaca fascicularis fascicularis*). *BMC Genomics* 16:222
2. Patel RP, Förster DW, Kitchener AC, et al (2016) Two species of Southeast Asian cats in the genus *Catopuma* with diverging histories: an island endemic forest specialist and a widespread habitat generalist. *Open Science* 3:160350
3. Roos C, Zinner D, Kubatko LS, et al (2011) Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. *BMC Evol Biol* 11:77
4. Liedigk R, Roos C, Brameier M, Zinner D (2014) Mitogenomics of the Old World monkey tribe Papionini. *BMC Evol Biol* 14:176
5. Fortes GG, Paijmans JLA (2015) Analysis of Whole Mitogenomes from Ancient Samples. *Methods Mol Biol* 1347:179–195
6. Lindgreen S (2012) AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res Notes* 5:337
7. Hahn C, Bachmann L, Chevreux B (2013) Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res.*
8. Schubert M, Ermini L, Der Sarkissian C, et al (2014) Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nat Protoc* 9:1056–1082
9. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–1760
10. McKenna A, Hanna M, Banks E, et al (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303
11. Xie Y, Wu G, Tang J, et al (2014) SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. *Bioinformatics* 30:1660–1666
12. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964
13. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinformatics* 10:421
14. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17:10–12
15. Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multiplatform graphical user interface

1
2
3
4 231 for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 27:221–224
5
6 232 16. Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA.
7 233 *Bioinformatics* 20:3252–3255
8
9 234 17. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements
10 235 in performance and usability. *Mol Biol Evol* 30:772–780
11
12 236 18. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
13 237 phylogenies. *Bioinformatics* 30:1312–1313
14

15 238
16
17 239
18
19 240
20
21 241
22
23 242
24

25
26 243 **Tables and Figures**

- 27
28 244 1. Table 1: List of mitogenomes assembled in this work that supplement pre-existing mitogenome
29 references already available in GenBank
30 245
31
32 246 2. Table 2: List of mitogenomes assembled in this work that currently have no complete
33 mitogenome reference available in GenBank
34 247
35
36 248 3. Figure 1: Geographical distribution of mitogenomes assembled in this work (62 mitogenomes with
37 known locality)
38 249
39
40 250 4. Figure 2: Phylogenetic tree of mitogenomes assembled in this work
41

42 251
43
44 252
45

46 253 **Additional files**

- 47
48 254 1. Additional file 1: DNA extraction of historical samples, library construction and primer information
49
50 255 2. Additional file 2: Sample information sheet of mitogenomes assembled in this work
51
52 256 3. Additional file 3: Alignment of mitogenomes assembled in this work
53
54 257 4. Additional file 4: Newick file for phylogenetic tree of mitogenomes assembled in this work
55
56

57 258
58
59 259
60

61
62
63
64
65

Table 1: List of mitogenomes assembled in this work that supplement pre-existing mitogenome references already available in GenBank

No	GenBank ID	Common name	Genus	Species	Assembly Size	Locality	Source	Sample date of collection	Data by
1	KY117537	<i>Hog deer</i>	<i>Axis</i>	<i>porcinus</i>	16402	CPH Zoo	ZM, KU	21/8/1912	FMS/FP
2	KY117538	<i>Pallas's squirrel</i>	<i>Callosciurus</i>	<i>erythraeus</i>	16656	Bangkok, Thailand	ZM, KU	25/5/1969	FMS/FP
3	KX265095	<i>Bay Cat</i>	<i>Catopuma</i>	<i>badia</i>	16,960	Sabah, Malaysia	National Museums Scotland	20/04/2000	PRP
4	KX224524	<i>Asiatic Golden Cat</i>	<i>Catopuma</i>	<i>temminckii</i>	16,960	Thailand	American Museum of National History, New York.	10/10/1927	PRP
5	KY117545	<i>Sumatran Rhino</i>	<i>Dicerorhinus</i>	<i>sumatrensis</i>	16466	Sumatra	Naturalis, Leiden, The Netherlands	1880	RM
6	KY117546	<i>Least pygmy squirrel</i>	<i>Exilisciurus</i>	<i>exilis</i>	16637	Indonesia	ROM	16/06/1993	FMS/FP
7	KY117548	<i>Hose's mongoose</i>	<i>Herpestes</i>	<i>javanicus</i>	16340	Java, Indonesia	ZM, KU	12/3/1947	FMS/FP
8	KY117550	<i>Three-striped ground Squirrel</i>	<i>Lariscus</i>	<i>indsignis</i>	16399	Maybe Malaysia	ZM, KU	unknown	FMS/FP
9	KY117592	<i>Black Crested Macaque</i>	<i>Macaca</i>	<i>nigra</i>	16558	captive	Gettorf Zoo, Germany	18/07/2000	CR
10	KY117593	<i>Northern Pig-tailed Macaque</i>	<i>Macaca</i>	<i>leonina</i>	16554	captive	Ludwig-Maximilians-University Munich, Germany	6/3/1995	CR
11	KY117594	<i>Southern pig-tailed macaque</i>	<i>Macaca</i>	<i>nemestrina</i>	16531	West Malaysia	National Museums Scotland, Edinburgh, UK	unknown	CR
12	KY117561	<i>Reeves's/Chinese Muntjak</i>	<i>Muntiacus</i>	<i>reevesi</i>	16354	Shin-Yi, Taiwan	Naturalis, Leiden, The Netherlands	1968	RM
13	KT288227	<i>Marbled Cat</i>	<i>Pardofelis</i>	<i>marmorata</i>	17,218	Sumatra	National archaeological museum of the Netherlands, Leiden	30/08/1930	PRP
14	KY117602	<i>Sumatra Surili</i>	<i>Presbytis</i>	<i>melalophos</i>	16,558	captive	Howletts Wild Animal Park, UK	23/7/1999	CR
15	KR135743	<i>Flat-headed Cat</i>	<i>Prionailurus</i>	<i>planiceps</i>	17,074	Sabah, Malaysia	Sabah wildlife department	25/04/2000	PRP
16	KY117580	<i>Malayan Field Rat</i>	<i>Rattus</i>	<i>tiomanicus</i>	16415	SPF Bidor, Perak, Malaysia	FRIM	12/2/2011	FMS/FP
17	KY117579	<i>Malayan Field Rat</i>	<i>Rattus</i>	<i>tiomanicus</i>	16312	Indonesia	ROM	01/06/1993	FMS/FP

18	KY117581	<i>Malayan Field Rat</i>	<i>Rattus</i>	<i>tiomanicus</i>	16305	Hutan Simpan Chikus, Tapah Perak, Malaysia	FRIM	13/1/2011	FMS/FP
19	KY117582	<i>Black Giant Squirrel</i>	<i>Ratufa</i>	<i>bicolor peninsulae</i>	16600	Bang Nara, Malakka, Thailand	ZM, KU	3/12/1932	FMS/FP
20	KY117574	<i>Javan Rhino</i>	<i>Rhinoceros</i>	<i>sondaicus</i>	16417	Java	Copenhagen Natural History Museum	unknown	RM
21	KY117575	<i>Javan Rusa</i>	<i>Rusa</i>	<i>timorensis</i>	16437	Toeloeng Agoeng, West Java	Naturalis, Leiden, The Netherlands	unknown	RM
22	KY117576	<i>Indian Sambar Deer</i>	<i>Rusa</i>	<i>unicolor dejeani</i>	16437	Mentawai, Indonesia	Naturalis, Leiden, The Netherlands	unknown	RM
23	KY117599	<i>Western Purple-faced Langur</i>	<i>Semnopithecus</i>	<i>vetulus</i>	16545	captive	Belfast Zoo, UK	9/11/1998	CR
24	KY117589	<i>Malayan Tapir</i>	<i>Tapirus</i>	<i>indicus</i>	16794	captive	Copenhagen Zoo	11/1/2015	FMS/FP
25	KY117598	<i>Silvered Langur</i>	<i>Trachypithecus</i>	<i>cristatus</i>	16551	North Sumatra	Bavarian State Collection Munich, Germany	1911	CR

ZM, KU: Zoological Musuem, University of Copenhagen; ROM: Royal Ontario Museum; FRIM: Forest Research Institute, Malaysia

Table 2: List of mitogenomes assembled in this work that currently have no complete mitogenome reference available in GenBank

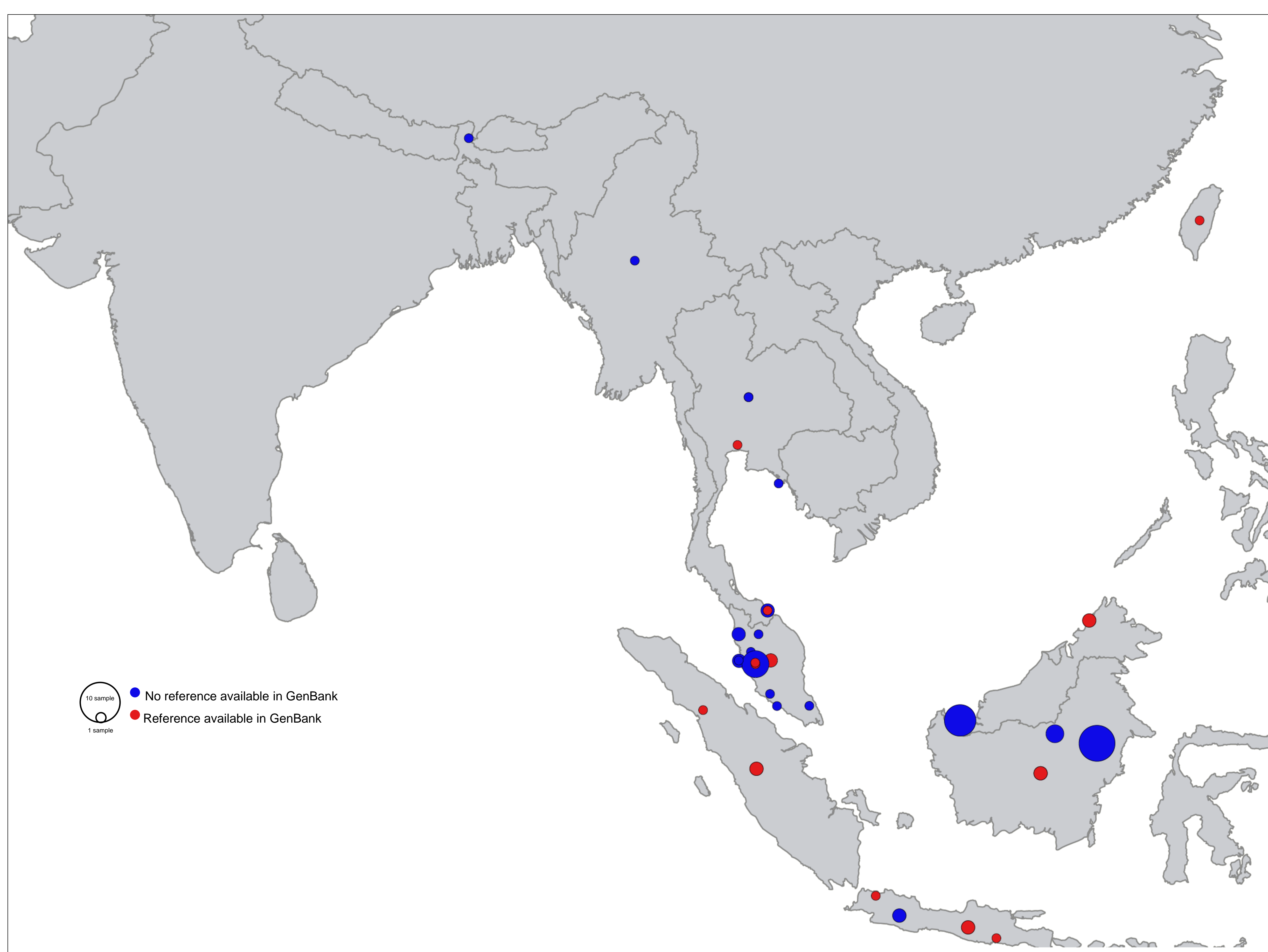
No	GenBank ID	Common name	Genus	Species	Assembly Size	Locality	Source	Sample date of collection	Data by
1	KY117536	<i>Asian Small-clawed Otter</i>	<i>Aonyx</i>	<i>cinereus</i>	16153	captive	Copenhagen Zoo	08/08/11	FMS/FP
2	KY117535	<i>Asian Small-clawed Otter</i>	<i>Aonyx</i>	<i>cinereus</i>	16153	Sarawak, Malaysia	British Musuem of Natural History, London Tierpark, Berlin	25/8/2010	FMS/FP
3	KY117560	<i>Binturong</i>	<i>Arctictis</i>	<i>binturong</i>	17,067	unknown		29/11/2010	PRP
4	KY117541	<i>Plantain Squirrel</i>	<i>Callosciurus</i>	<i>notatus</i>	16582	Hutan Bidor, Perak, Malaysia	FRIM	11/2/2011	FMS/FP
5	KY117542	<i>Plantain squirrel</i>	<i>Callosciurus</i>	<i>notatus</i>	16599	East Kalimantan, Indonesia	ROM	03/06/1993	FMS/FP
6	KY117543	<i>Prevost's squirrel</i>	<i>Callosciurus</i>	<i>prevostii</i>	16674	East Kalimantan, Indonesia	ROM	15/06/1993	FMS/FP
7	KY117540	<i>Variable squirrel</i>	<i>Callosciurus</i>	<i>finlaysonii frandseni</i>	15747	Koh Chang	ZM, KU	14/1/1900	FMS/FP
8	KY117539	<i>Variable squirrel</i>	<i>Callosciurus</i>	<i>finlaysonii</i>	16489	Central Siam	ZM, KU	2/2/1928	FMS/FP
9	KY117544	<i>Sunda Otter Civet</i>	<i>Cynogale</i>	<i>bennetti</i>	15784	Borneo	British Musuem of Natural History, London	25/8/2010	FMS/FP
10	KY117547	<i>Hose's mongoose</i>	<i>Herpestes</i>	<i>hosei</i>	16819	Ceylon	ZM, KU	1/7/1931	FMS/FP
11	KY117549	<i>Thick-spined porcupine</i>	<i>Hystrix</i>	<i>crassipinis</i>	15778	CPH Zoo	ZM, KU	19/1/1874	FMS/FP
12	KY117552	<i>Long-tailed Giant Rat</i>	<i>Leopoldamys</i>	<i>sabanus</i>	15973	G. Telapak Buruk, Negeri Sembilan, Malaysia	FRIM	24/2/2010	FMS/FP
13	KY117553	<i>Long-tailed Giant Rat</i>	<i>Leopoldamys</i>	<i>sabanus</i>	15972	Teluk Segadas, P. Pangkor, Perak, Malaysia	FRIM	19/3/2010	FMS/FP
14	KY117554	<i>Long-tailed Giant Rat</i>	<i>Leopoldamys</i>	<i>sabanus</i>	15974	Hutan Simpan Temengor, Gerik Perak, Malaysia	FRIM	23/1/2014	FMS/FP
15	KY117555	<i>Long-tailed Giant Rat</i>	<i>Leopoldamys</i>	<i>sabanus</i>	15972	Hutan Simpan Lenggong, Kluang, Johor, Malaysia	FRIM	19/2/2014	FMS/FP
16	KY117551	<i>Long-tailed giant rat</i>	<i>Leopoldamys</i>	<i>sabanus</i>	15974	Malaysia	ROM	28/05/1993	FMS/FP
17	KY117556	<i>Hairy-nosed otter</i>	<i>Lutra</i>	<i>sumatrana</i>	16580	Bang Nara, Thailand	ZM, KU	1/4/1939	FMS/FP

18	KY117557	<i>Smooth-coated otter</i>	<i>Lutrogale</i>	<i>perspicillata</i>	16042	Malacca, Malaysia	British Museum of Natural History, London	25/8/2010	FMS/FP
19	KY117558	<i>Smooth-coated otter</i>	<i>Lutrogale</i>	<i>perspicillata</i>	16041	Bang Nara, Thailand	ZM, KU	24/1/1933	FMS/FP
20	KY117591	<i>Moor macaque</i>	<i>Macaca</i>	<i>maura</i>	16563	captive	Hannover Zoo, Germany	20/8/1998	CR
21	KY117564	<i>Rajah/Brown Spiny Rat</i>	<i>Maxomys</i>	<i>rajah</i>	16200	Indonesia	ROM	06/06/1993	FMS/FP
22	KY117562	<i>Rajah/Brown Spiny Rat</i>	<i>Maxomys</i>	<i>rajah</i>	16296	Teluk Segadas, P. Pangkor, Perak, Malaysia	FRIM	19/3/2010	FMS/FP
23	KY117563	<i>Rajah/Brown Spiny Rat</i>	<i>Maxomys</i>	<i>rajah</i>	16296	Pasir Bogak, P. Pangkor, Perak, Malaysia	FRIM	18/3/2010	FMS/FP
24	KY117567	<i>Red Spiny Rat</i>	<i>Maxomys</i>	<i>surifer</i>	16286	50 ha, Pasoh, Negeri Sembilan, Malaysia	FRIM	12/6/2008	FMS/FP
25	KY117566	<i>Red Spiny Rat</i>	<i>Maxomys</i>	<i>surifer</i>	16290	Indonesia	ROM	21/05/1993	FMS/FP
26	KY117565	<i>Red Spiny Rat</i>	<i>Maxomys</i>	<i>surifer</i>	16286	Malaysia	ROM	17/05/2013	FMS/FP
27	KY117570	<i>Whitehead's Spiny Rat</i>	<i>Maxomys</i>	<i>whiteheadi</i>	16316	Hutan Simpan Bikam, Perak, Malaysia	FRIM	12/2/2011	FMS/FP
28	KY117571	<i>Whitehead's Spiny Rat</i>	<i>Maxomys</i>	<i>whiteheadi</i>	16316	Keruing Trail, FRIM, Kepong, Selangor, Malaysia	FRIM	13/3/2013	FMS/FP
29	KY117568	<i>Whitehead's Spiny Rat</i>	<i>Maxomys</i>	<i>whiteheadi</i>	16287	Hutan Simpan Bikam, Perak, Malaysia	FRIM	12/2/2011	FMS/FP
30	KY117569	<i>Whitehead's Spiny Rat</i>	<i>Maxomys</i>	<i>whiteheadi</i>	16429	Bukit Tapah, Perak, Malaysia	FRIM	23/3/2011	FMS/FP
31	KY052142	<i>Indian muntjac</i>	<i>Muntiacus</i>	<i>muntjak</i>	16354	West Java	Vienna NHM	1858	RM
32	KY117559	<i>Bornean yellow muntjac</i>	<i>Muntiacus</i>	<i>atherodes</i>	16354	Koemai, West Borneo	Bonn NHM	1938	RM
33	KY052092	<i>Indian Red Muntjac</i>	<i>Muntiacus</i>	<i>vaginalis</i>	16354	Gangtok, South Sikkim, India	Berlin NHM	unknown	RM
34	KY117573	<i>Dark-tailed Tree Rat</i>	<i>Niviventer</i>	<i>cremoriventer</i>	16322	Track 5 (G.Inas), Kedah, Malaysia	FRIM	5/11/2009	FMS/FP
35	KY117572	<i>Dark-tailed tree rat</i>	<i>Niviventer</i>	<i>cremoriventer</i>	16234	Malaysia	ROM	17/05/2013	FMS/FP

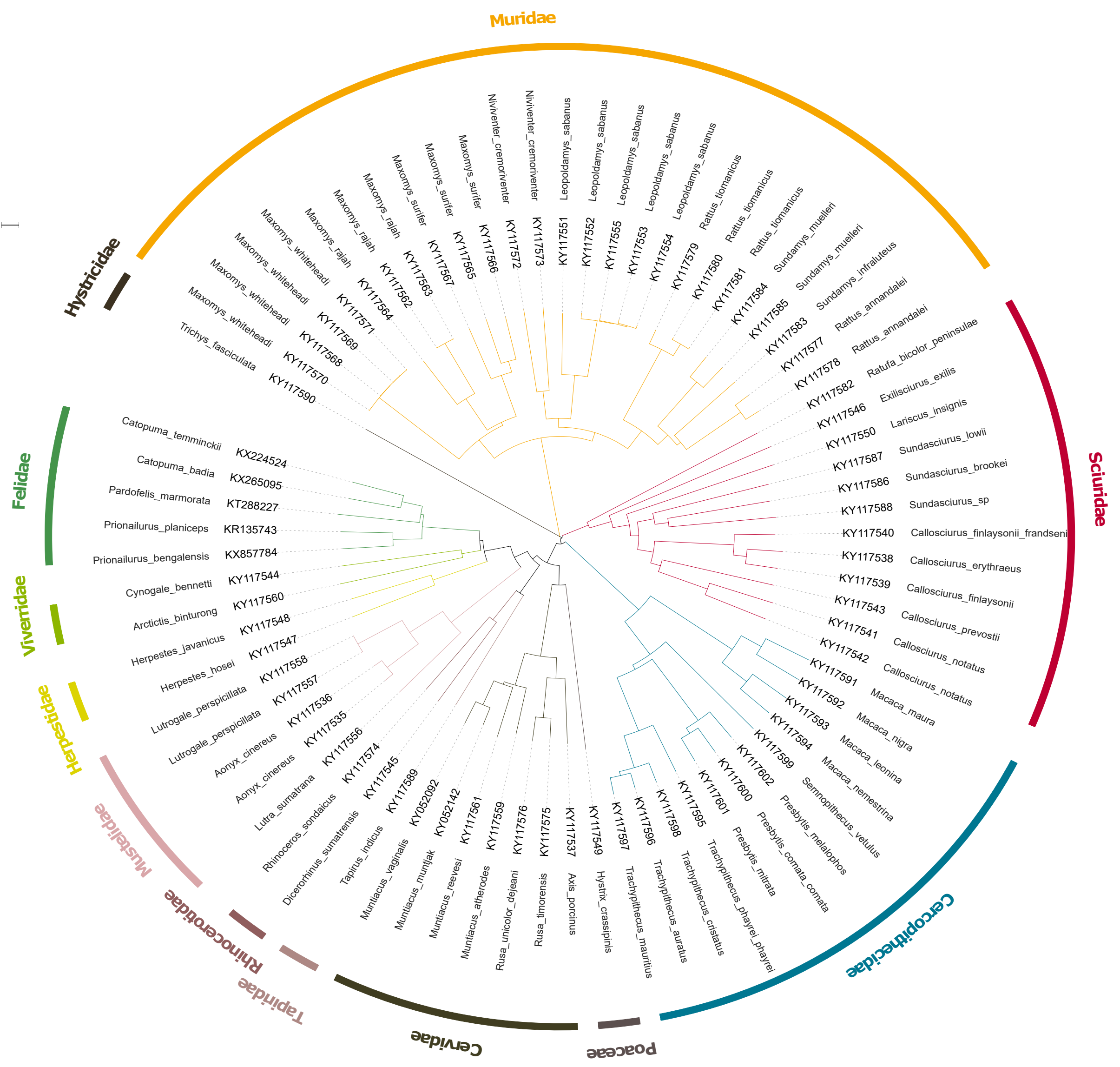
36	KY117600	<i>Grizzled Leaf Monkey</i>	<i>Presbytis</i>	<i>comata comata</i>	16551	captive	Howletts Wild Animal Park, UK	23/12/1999	CR
37	KY117601	<i>Mitred Leaf Monkey</i>	<i>Presbytis</i>	<i>mitrata</i>	16,555	captive	Howletts Wild Animal Park, UK	12/11/1998	CR
38	KX857784	<i>Leopard cat</i>	<i>Prionailurus</i>	<i>bengalensis</i>	16,989	Thailand	American Museum of Natural History, New York.	25/02/1924	PRP
39	KY117578	<i>Annandale's Sundaic Rat</i>	<i>Rattus</i>	<i>annandalei</i>	16297	Hutan Simpan Bikam, Perak, Malaysia	FRIM	12/2/2011	FMS/FP
40	KY117577	<i>Annandale's Sundaic Rat</i>	<i>Rattus</i>	<i>annandalei</i>	16301	Hutan Simpan Bikam, Perak, Malaysia	FRIM	11/2/2011	FMS/FP
41	KY117583	<i>Mountain giant Sunda rat</i>	<i>Sundamys</i>	<i>infraluteus</i>	16297	Malaysia	ROM	18/05/2013	FMS/FP
42	KY117585	<i>Müller's Giant Sunda Rat</i>	<i>Sundamys</i>	<i>meulleri</i>	16326	Track 1 (G.Inas), Kedah, Malaysia	FRIM	5/11/2009	FMS/FP
43	KY117584	<i>Müller's giant Sunda rat</i>	<i>Sundamys</i>	<i>meulleri</i>	16304	Malaysia	ROM	01/06/2013	FMS/FP
44	KY117586	<i>Brooke's squirrel</i>	<i>Sundasciurus</i>	<i>brookei</i>	16417	East Kalimantan, Indonesia	ROM	13/06/1993	FMS/FP
45	KY117587	<i>Low's squirrel</i>	<i>Sundasciurus</i>	<i>lowii</i>	16307	East Kalimantan, Indonesia	ROM	06/06/1993	FMS/FP
46	KY117588	<i>name pending</i>	<i>Sundasciurus</i>	<i>sp</i>	16458	East Kalimantan, Indonesia	ROM	21/06/1993	FMS/FP
47	KY117595	<i>Phayre's Langur</i>	<i>Trachypithecus</i>	<i>phayrei phayrei</i>	16548	South West Myanmar	Natural History Museum Berlin, Germany	unknown	CR
48	KY117596	<i>East Javan Ebony Langur</i>	<i>Trachypithecus</i>	<i>auratus</i>	16552	captive	Bristol Zoo, UK	26/10/2010	CR
49	KY117597	<i>West Javan Ebony Langur</i>	<i>Trachypithecus</i>	<i>mauritus</i>	16554	West Java	Naturalis Leiden; Netherlands	unknown	CR
50	KY117590	<i>Long-tailed porcupine</i>	<i>Trichys</i>	<i>fasciculata</i>	16328	Borneo	ZM, KU	5/10/1894	FMS/FP

ZM, KU: Zoological Museum, University of Copenhagen; ROM: Royal Ontario Museum; FRIM: Forest Research Institute, Malaysia; NHM: Natural History

/ Musuem



Tree scale: 0.01

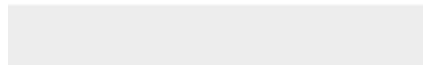




Click here to access/download

Supplementary Material

1 Data information_Sallehetal.docx





Click here to access/download
Supplementary Material
1 README file_Salleh.txt

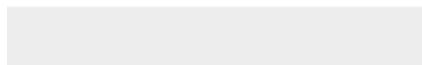




[Click here to access/download](#)

Supplementary Material

Additional file 1_Sallehetal.docx





Click here to access/download
Supplementary Material
Additional file 2_Sallehetal.xlsx





Click here to access/download
Supplementary Material
Additional file 3_Sallehetal.fa





Click here to access/download
Supplementary Material
Additional file 4 newick.txt

