

Original article

Field-deployable genetic methods for subspecies identification of long-tailed macaques (*Macaca fascicularis*) based on Oxford Nanopore technologies

Juthamas Nimnual¹, Vorthon Sawaswong², Suthida Visedthorn³, Ariya Khamwut³, Taratorn Kemthong⁴, Suchinda Malaivijitnond⁴, Sunchai Payungporn^{3,*}

¹Program in Medical Sciences, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

²Department of Biochemistry, Faculty of Science, Mahidol University, Bangkok, Thailand

³Center of Excellence in Systems Microbiology, Department of Biochemistry, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

⁴Department of Biology, Faculty of Science, Chulalongkorn University, Bangkok, Thailand

Abstract

Background: Long-tailed macaque (*Macaca fascicularis*) is an ecologically and biomedically important primate widely distributed throughout Southeast Asia. Accurate subspecies identification is essential for conservation, ecological monitoring, and biomedical research. Morphological classification is often unreliable in hybrids.

Objective: To develop and evaluate a field-deployable genetic workflow for subspecies identification of *M. fascicularis* using fecal swab samples and Oxford Nanopore Technologies (ONT).

Methods: Fecal swab samples were collected from 30 individuals representing *M. fascicularis fascicularis* (n = 10), *M. fascicularis aurea* (n = 10), and hybrids (*Mff* × *Mfa*; n = 10). DNA from NAPSeq-preserved swabs was PCR-amplified for the mitochondrial hypervariable segment I (HVSI) and the sex-determining region Y (SRY) gene, followed by ONT sequencing and phylogenetic analysis.

Results: All samples yielded adequate DNA concentration (37.6 ± 17.7 ng/μL) and purity ($A_{260}/A_{280} = 1.90 \pm 0.12$). HVSI (881 bp) and SRY (846 bp) target genes were successfully amplified. ONT sequencing produced 560,171 raw reads (HVSI) and 220,740 raw reads (SRY). Phylogenetic analysis based on HVSI and SRY determined maternal and paternal lineages, respectively, enabling identification of subspecies and hybrid ancestry. Subspecies classification was concordant between morphology-based and ONT-based analyses for all samples, demonstrating the reliability and accuracy of the genetic workflow.

Conclusion: ONT-based sequencing of DNA extracted from fecal samples provides a practical and reliable approach for field-deployable subspecies identification in *M. fascicularis*, which would be applicable for ecological surveillance and quality control of non-human primate models in biomedical research.

Keywords: long-tailed macaques, *Macaca fascicularis*, Oxford Nanopore Technologies, subspecies identification

***Correspondence to:** Sunchai Payungporn, Center of Excellence in Systems Microbiology, Department of Biochemistry, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand
E-mail: juthamas.cu@gmail.com, sp.medbiochemcu@gmail.com

Received: 1 May 2026

Revised: 20 June 2026

Accepted: 30 June 2026

The genus *Macaca* comprises one of the most evolutionarily successful and geographically widespread groups of Old-World primates, occupying diverse habitats across East, Southeast, and South Asia.^(1, 2) Among these, the long-tailed macaque (*Macaca fascicularis*) is the most commonly encountered non-human primate (NHP) species in Thailand, with two major subspecies recognized: *M. fascicularis* (common long-tailed macaque) and *M. fascicularis aurea* (Burmese long-tailed macaque), as well as hybrid populations arising from interbreeding between these subspecies and with *M. mulatta*.^(3, 4) The biological importance of *M. fascicularis* is underscored by its close genetic similarity to humans (approximately 92.8%–95%), which makes it an indispensable animal model in biomedical research, including studies of infectious diseases, vaccine development, and drug safety evaluation.^(5–7) Accurate species and subspecies identification is fundamental for ecological monitoring, population management, and selection of appropriate animal models in biomedical research. Traditionally, macaque classification has relied on morphological characteristics including body size, pelage coloration, relative tail length, and craniofacial features.^(8, 9) However, phenotypic plasticity, interspecies hybridization, and overlapping morphological traits frequently render visual identification unreliable, particularly in hybrid zones where individuals may display intermediate phenotypes that obscure their true genetic ancestry.^(10, 11) Furthermore, in field-based studies, researchers commonly observe macaques from a distance to minimize animal stress, limiting accurate morphological assessment. These constraints necessitate the adoption of molecular genetic approaches for objective and reproducible species identification.

Mitochondrial DNA (mtDNA) markers, particularly the hypervariable segment I (HVSI) of the D-loop region, are widely employed for maternal lineage analysis and species discrimination in macaques owing to their high mutation rate, maternal inheritance, and abundance in cells.⁽¹²⁾ Phylogenetic analyses based on HVSI have previously identified five distinct maternal subclades within *M. fascicularis* across the Indochinese and Sundaic regions, including specific populations in Thailand.⁽¹²⁾ Complementarily, Y-chromosomal markers such as the sex-determining region Y (SRY) gene provide a paternal lineage perspective and enable discrimination of male-

mediated gene flow patterns across subspecies boundaries.^(13, 14) The combined use of these two markers allows simultaneous assessment of maternal and paternal lineages, which is particularly informative for identifying hybrid individuals and characterizing the genetic composition of admixed populations.

Oxford Nanopore Technologies (ONT) is a third-generation sequencing platform that offers several advantages for field-based genetic studies, including portability, real-time data acquisition, long-read sequencing capability, and compatibility with low-input DNA samples.⁽¹⁵⁾ The MinION device is particularly suited for deployment outside traditional laboratory environments. Recent studies have demonstrated the utility of ONT for rapid species identification and phylogenetic analysis in wildlife forensics and conservation genetics.⁽¹⁶⁾ However, reliable non-invasive molecular approaches for subspecies identification and hybrid detection in *M. fascicularis* remain limited, particularly under field conditions where animal capture and invasive sampling are impractical. Furthermore, the application of ONT sequencing to fecal swab samples for distinguishing *Mff*, *Mfa*, and their hybrids has not been systematically evaluated. Therefore, this study aimed to address these limitations. genetic workflow for subspecies identification of *M. fascicularis* using fecal swab samples and ONT sequencing with practical applications in ecological surveillance, conservation management, and biomedical research.

Materials and methods

Overview procedure

This study was conducted under Animal Use Protocol No. 2575009, approved by the Institutional Animal Care and Use Committee (IACUC) of National Primate Research Center of Thailand, Chulalongkorn University (NPRCT-CU). Morphological subspecies identification was performed by experienced veterinarians based on established taxonomic criteria for the genus *Macaca*.^(8, 9) The workflow comprised five steps (**Figure 1**): fecal swab collection, DNA extraction, PCR amplification, long read nanopore sequencing, and phylogenetic analysis.

Sample collection and DNA extraction

Fecal swab samples were collected using sterile cotton swabs and preserved in 2 mL of NAPSeq buffer (Bioentist, Thailand) for DNA stabilization prior to storage at -20°C . Fecal swab samples were

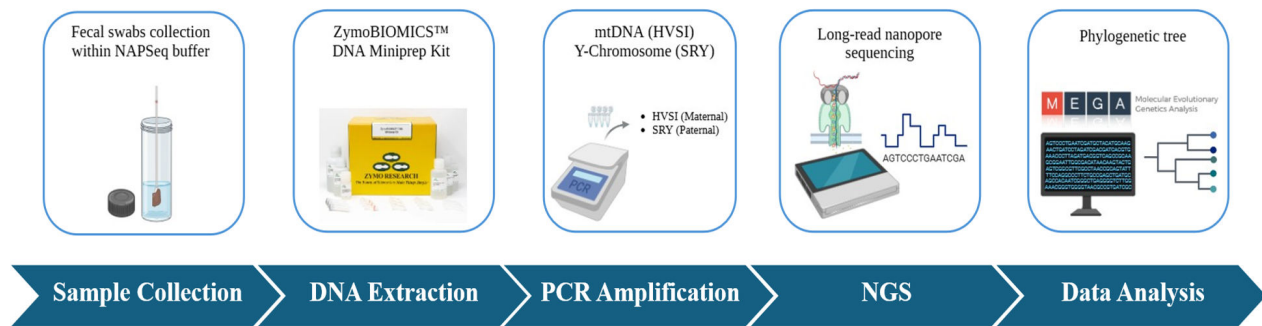


Figure 1. Overview of genetic workflow for subspecies identification.

Table 1. Primer sequences in this study.

Target	Primer	Sequence (5' → 3')	Amplicon size (bp)
HVSI	HVS-F	<u>TTTCTGTTGGTGCTGATATTGCCGCCC</u> ACTCAGCCAATTCCTGTTCT	881 bp
	HVS-R	ACTTGCCCTGTCGCTCTATCTTC <u>CCCCGTGATCCATCGAGATGTCTT</u>	
SRY	SW2-F	<u>TTTCTGTTGGTGCTGATATTGC</u> CTTGAGAATACATTGTCAGGG	846 bp
	SW3B-R	ACTTGCCCTGTCGCTCTATCTTCAGGTCTTTGTAGCCAATGTTACCCG	

Underlined sequences represent ONT adapter tails. HVSI, hypervariable segment I; SRY, sex-determining region Y.

extracted using the ZymoBIOMICS™ DNA Miniprep Kit (Zymo Research, USA). Briefly, each sample was vortexed for 30 secs, after which the cotton swab was transferred into a ZR BashingBead™ Lysis Tube. The samples were centrifuged at 4,000 rpm for 15 min, and the NAPSeq buffer was discarded. Next, 750 µL of lysis buffer was added to the same tube, and the cotton swab was reinserted and vortexed for 30 secs to ensure thorough cell lysis. The lysate together with the cotton swab was then homogenized in the ZR BashingBead™ Lysis Tube at 50 Hz for 3 min. DNA extraction was subsequently completed according to the manufacturer's protocol. The concentration of the extracted DNA was measured using a NanoPhotometer® C40 (Implen, Germany). The DNA samples were stored at –20°C until further use.

PCR amplification of HVSI and SRY

PCR amplification was performed in a final reaction volume of 10 µL containing 5 µL of KOD One™ PCR Master Mix Blue (Toyobo Life Science, Japan), 0.3 µL of each forward and reverse primer (10 µM), 3.4 µL of DEPC-treated water, and 1 µL of 10-fold diluted DNA template. Two target genes including HVSI region and SRY were amplified for maternal and paternal lineages classification, respectively. Both primer pairs were modified from previously published sequences by adding tailed primers incorporating ONT adapter sequences to

facilitate direct nanopore library preparation (**Table 1**).^(12, 13) The PCR amplification was performed by Proflex PCR System (Thermo Fisher Scientific, USA). The first round PCR cycling conditions consisted of an initial denaturation at 98°C for 2 min, followed by 40 cycles of amplification denaturation at 98°C for 10 sec, annealing at 60°C for 10 sec (HVSI) or 56°C for 10 secs (SRY), and extension at 68°C for 10 sec. A final extension step was performed at 68°C for 5 min. A second round of PCR was performed to attach barcodes using the PCR Barcoding Expansion 1-96 kit. The reaction began with an initial denaturation at 98°C for 2 min, followed by 5 cycles of amplification consisting of denaturation at 98°C for 10 secs, annealing at 60°C for 10 secs, and extension at 68°C for 10 secs. A final extension step was carried out at 68°C for 5 min. PCR products were examined by electrophoresis on a 1% agarose gel stained with RedSafe (iNtRON Biotechnology, South Korea) and visualized using a ChemiDoc Imaging System (Bio-Rad Laboratories, USA).

Library preparation and nanopore sequencing

PCR products used for species identification were purified using the HiYield™ Gel/PCR DNA Fragments Extraction Kit (RBC Bioscience, Taiwan) according to the manufacturer's protocol to remove residual primers, nucleotides, and enzymes. The concentration of the purified DNA was subsequently quantified using a Qubit Fluorometer (Thermo Fisher Scientific, USA).

Table 2. Accession no. of reference sequences in this study.

Species of macaques	Accession no. for HVSI	Accession no. for SRY
<i>M. sylvanus</i>	NC002764	AF284326
<i>M. fascicularis fascicularis</i>	LC536646	LC225413
<i>M. fascicularis aurea</i>	OQ199573	LC225411
<i>M. mulatta</i>	NC005943	LC093268
<i>M. nemestrina</i>	NC026976	NC092146
<i>M. leonina</i>	NC027604	KT356228
<i>Mff</i> × <i>Mfa</i>	LC093191	N/A

Purified DNA samples were then pooled to a final concentration of 1 µg for multiplexing and size-selected by clean-up with 0.5X Agencourt AMPure XP beads (Beckman Coulter, USA). Library preparation was performed using the Ligation Sequencing Kit SQK-LSK114 following the manufacturer's instructions to repair and ligate the DNA ends for library construction. Sequencing was conducted on a MinION™ Mk1C device using an R10.4.1 flow cell, with data acquisition performed through MinKNOW software (Oxford Nanopore Technologies, UK).

Data processing and phylogenetic tree analysis

Basecalling of POD5 files was performed using Dorado v1.4 (model: dna_r10.4.1_e8.2_400bps_sup @v5.2.0) to generate FASTQ sequences. The FASTQ files were subsequently demultiplexed and adapter-trimmed using MinKNOW v25.05.14. The trimmed reads were mapped to reference mitochondrial DNA (mtDNA) and Y-chromosomal sequences using Minimap2 v2.31 (-ax map-ont). Consensus sequences were first generated using SAMtools v1.23.1 (samtools consensus) and subsequently polished using Medaka v2.2.2 (medaka_consensus -m r1041_e82_400bps_sup_v5.0.0) to improve sequence accuracy. The resulting consensus sequences were used for downstream phylogenetic analyses. For phylogenetic analysis, reference sequences were retrieved from GenBank and aligned with the consensus sequences obtained in this study (**Table 2**). Multiple sequence alignment was performed using the ClustalW algorithm implemented in Molecular Evolutionary Genetics Analysis using MEGA X. The best-fit nucleotide substitution model was evaluated based on the Bayesian Information Criterion (BIC). Phylogenetic relationships were inferred using the Maximum Likelihood (ML) method, with branch support assessed through 1,000 bootstrap replicates. The resulting phylogenetic trees were visualized and annotated using the Interactive Tree of Life (iTOL) v6 platform to facilitate graphical representation and interpretation.

Results

DNA quality and quantity

DNA was successfully extracted from all 30 fecal swab samples preserved in NAPSeq buffer. The mean DNA concentration across all samples was 37.6 ± 17.7 ng/µL (range: 10.0–87.1 ng/µL), with comparable yields across the three groups. DNA purity assessed by A260/A280 absorbance ratios ranged from 1.6 to 2.2, with a mean of 1.90 ± 0.12 , and 24 out of 30 samples exhibited ratios within the optimal range of 1.8–2.0 (**Table 3**). These results demonstrate that NAPSeq buffer effectively preserves fecal swab DNA at room temperature and that the ZymoBIOMICS™ DNA Miniprep Kit provides sufficient quality and quantity of extracted DNA for downstream molecular analysis.

PCR amplification of HVSI and SRY genes

All 30 fecal swab samples yielded sufficient DNA for PCR amplification. Successful amplification of the HVSI amplicon (881 bp) was confirmed by 1% agarose gel electrophoresis (**Figure 2A**). The SRY amplicon (846 bp) was detected exclusively in the 15 male samples across all three groups, consistent with the Y-chromosomal and male-specific marker (**Figure 2B**). No amplification was observed in the no template control (N) for either target, suggesting the absence of contamination. These results confirm the feasibility of using fecal swab DNA for multi-target amplification of both mitochondrial and Y-chromosomal loci.

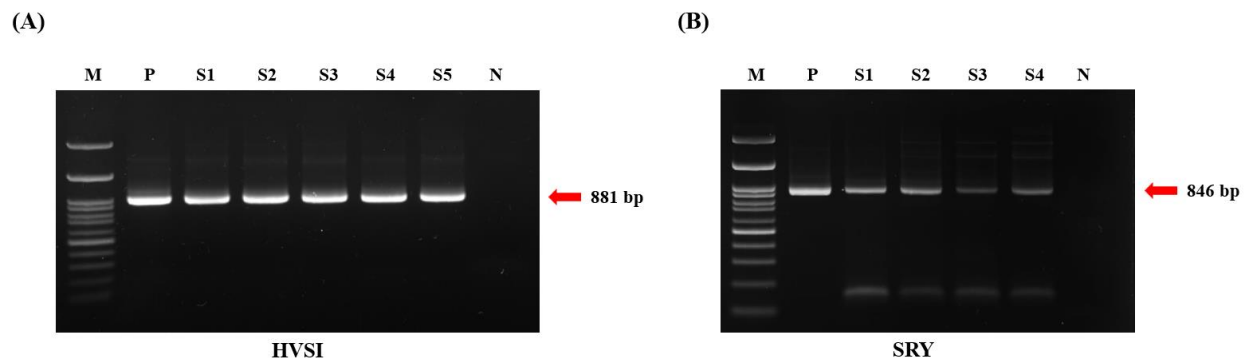
Sequencing output and data processing

For HVSI sequencing, a total of 560,171 raw reads were generated from all 30 samples, of which 304,995 reads were retained following quality filtering at a minimum Phred score of Q15. The mean percent identity of consensus sequences against reference sequences was $94.06 \pm 3.68\%$. As for SRY sequencing, 220,740 raw reads were generated from 15 male samples, with 43,751 reads retained after Q15 filtering. The mean percent identity of SRY consensus sequences was $99.94 \pm 0.08\%$.

Table 3. Summary of sample collection, DNA concentration, DNA purity, geographical information and subspecies identification based on morphological assessment and ONT-based phylogenetic analysis for all fecal swab samples.

Monkey ID	Sex	DNA Concentration (ng/μL)	DNA purity (A260/A280)	Location	Province	GPS (North, East)	Subspecies identification		
							based on morpholog	based on ONT	
								Maternal lineage	Paternal lineage
KRI005	M	34.0	1.9				<i>Mff</i> x <i>Mfa</i>	<i>Mff</i> x <i>Mfa</i>	<i>Mff</i>
KRI006	M	24.1	1.9				<i>Mff</i> x <i>Mfa</i>	<i>Mff</i> x <i>Mfa</i>	<i>Mff</i>
KRI012	M	61.6	2.0				<i>Mff</i> x <i>Mfa</i>	<i>Mff</i> x <i>Mfa</i>	<i>Mff</i>
KRI014	M	73.6	1.9				<i>Mff</i> x <i>Mfa</i>	<i>Mff</i> x <i>Mfa</i>	<i>Mfa</i>
KRI023	M	26.1	2.0	Kho Ram Island (KRI)	Prachuap Khiri Khan	12°14'33.0" 100°00'31.0"	<i>Mff</i> x <i>Mfa</i>	<i>Mff</i> x <i>Mfa</i>	<i>Mff</i>
KRI001	F	48.5	1.9				<i>Mff</i> x <i>Mfa</i>	<i>Mff</i> x <i>Mfa</i>	N/A
KRI002	F	40.8	1.9				<i>Mff</i> x <i>Mfa</i>	<i>Mff</i> x <i>Mfa</i>	N/A
KRI008	F	20.2	1.8				<i>Mff</i> x <i>Mfa</i>	<i>Mff</i> x <i>Mfa</i>	N/A
KRI010	F	38.8	1.9				<i>Mff</i> x <i>Mfa</i>	<i>Mff</i> x <i>Mfa</i>	N/A
KRI011	F	14.1	2.0				<i>Mff</i> x <i>Mfa</i>	<i>Mff</i> x <i>Mfa</i>	N/A
TPK004	M	43.5	1.9				<i>Mfa</i>	<i>Mfa</i>	<i>Mfa</i>
TPK010	M	43.9	1.9				<i>Mfa</i>	<i>Mfa</i>	<i>Mfa</i>
TPK013	M	27.4	2.0				<i>Mfa</i>	<i>Mfa</i>	<i>Mfa</i>
TPK015	M	18.7	2.1				<i>Mfa</i>	<i>Mfa</i>	<i>Mfa</i>
TPK020	M	45.0	1.9	Tham Pra Khayang (TPK)	Ranong Province	10°19'35.8" 98°45'53.1"	<i>Mfa</i>	<i>Mfa</i>	<i>Mfa</i>
TPK003	F	34.5	1.7				<i>Mfa</i>	<i>Mfa</i>	N/A
TPK008	F	37.9	1.7				<i>Mfa</i>	<i>Mfa</i>	N/A
TPK009	F	50.0	1.9				<i>Mfa</i>	<i>Mfa</i>	N/A
TPK018	F	26.6	1.6				<i>Mfa</i>	<i>Mfa</i>	N/A
TPK027	F	31.4	1.7				<i>Mfa</i>	<i>Mfa</i>	N/A
PPT165	M	87.1	1.9				<i>Mff</i>	<i>Mff</i>	<i>Mff</i>
PPT185	M	71.3	1.9				<i>Mff</i>	<i>Mff</i>	<i>Mff</i>
PPT195	M	29.9	2.0				<i>Mff</i>	<i>Mff</i>	<i>Mff</i>
PPT151	M	66.30	2.0				<i>Mff</i>	<i>Mff</i>	<i>Mff</i>
PPTC216	M	31.6	1.9	NPRCT-CU (PPT/PPTC)	Saraburi Province	14°31'32.9" 101°01'02.5"	<i>Mff</i>	<i>Mff</i>	<i>Mff</i>
PPTC125	F	20.3	1.9				<i>Mff</i>	<i>Mff</i>	N/A
PPTC128	F	26.7	1.9				<i>Mff</i>	<i>Mff</i>	N/A
PPTC165	F	35.6	2.0				<i>Mff</i>	<i>Mff</i>	N/A
PPTC199	F	30.4	1.9				<i>Mff</i>	<i>Mff</i>	N/A
PPTC219	F	45.7	1.9				<i>Mff</i>	<i>Mff</i>	N/A

M, Male; F, Female; N/A, not available

**Figure 2.** Agarose gel electrophoresis of PCR amplification products from representative fecal swab samples. (A) HVSI amplification showing the expected 881 bp PCR products (S1–S5). (B) SRY amplification showing the expected 846 bp in male samples (S1–S4). M = 100 bp DNA ladder, P = positive control, N = no template control.

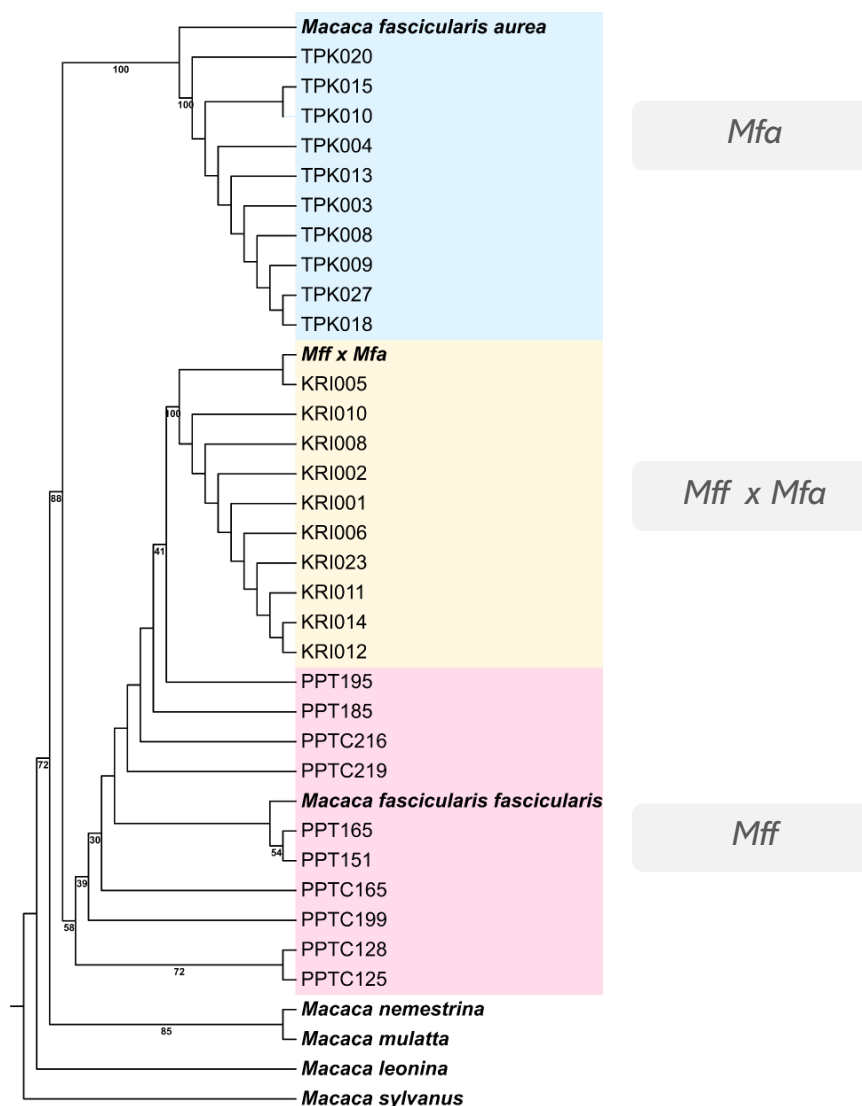


Figure 3. Maximum Likelihood phylogenetic tree based on mitochondrial HVSI sequences. Colored highlights indicate lineage groups: blue = *Mfa* (TPK); yellow = *Mff* × *Mfa* (KRI); pink = *Mff* (PPT/PPTC). *Macaca sylvanus* was used as the outgroup.

Subspecies identification by HVSI phylogenetic analysis

Phylogenetic analysis of HVSI sequences from all fecal swab samples using the Maximum Likelihood method resolved three distinct lineage clusters with clear subspecies-level separation (**Figure 3**). All *Mfa* samples formed a well-supported clade (bootstrap = 100) clustering with the *Mfa* reference sequence at the top of the tree, highlighted in blue. All hybrid samples clustered within a distinct clade together with previously documented *Mff* × *Mfa* hybrid reference sequences from the same geographic contact zone (highlighted in yellow), indicating a shared maternal lineage consistent with matrilineal descent from a

locally established hybrid population.⁽¹²⁾ All *Mff* samples were resolved within the *Mff* reference clade, highlighted in pink, grouping alongside with the *Mff* reference sequence. Outgroup taxa (*M. nemestrina*, *M. mulatta*, *M. leonina*, and *M. sylvanus*) were correctly resolved as basal lineages outside the *M. fascicularis* group, with *M. sylvanus* occupying the most distal outgroup position. These results demonstrate that HVSI-based nanopore sequencing can reliably resolve maternal lineage structure among *Mff*, *Mfa*, and hybrid individuals from fecal swab-derived DNA.

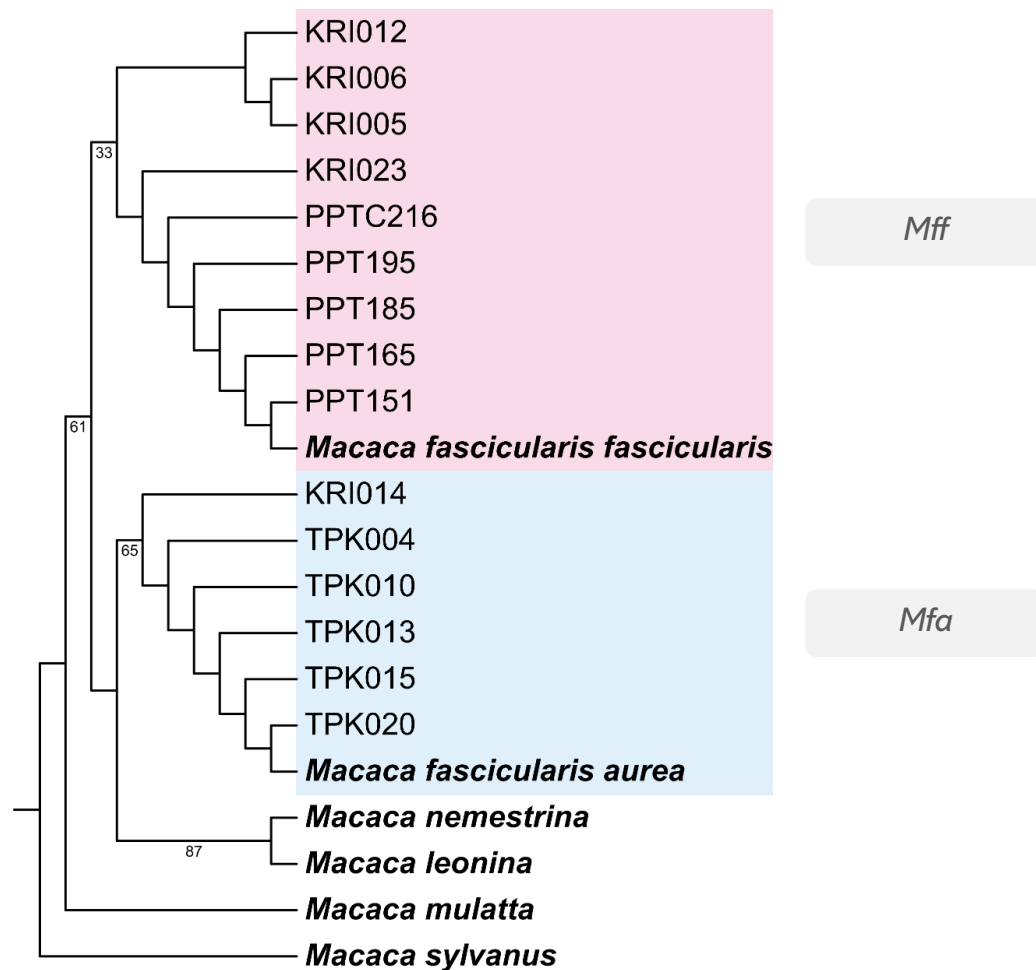


Figure 4. Maximum Likelihood phylogenetic tree based on Y-chromosomal SRY sequences. Colored highlights indicate lineage groups: blue = *Mfa* (TPK) with hybrid KRI014; pink = *Mff* (PPT/PPTC) with hybrid KRI samples. *Macaca sylvanus* was used as the outgroup.

Subspecies identification by SRY phylogenetic analysis

Phylogenetic analysis of SRY sequences from 15 male samples using the Maximum Likelihood method discriminated two major clusters corresponding to *Mfa* and *Mff* paternal lineages (**Figure 4**). All *Mfa* male samples (TPK004, TPK010, TPK013, TPK015, TPK020), together with hybrid individual (KRI014), formed a clade with the *Mfa* reference sequence (highlighted in blue). All *Mff* male samples (PPT151, PPT165, PPT185, PPT195, PPTC216) clustered within the *Mff* reference clade (highlighted in pink), closely related to the *Mff* reference sequence. Notably, *M. sylvanus*, *M. mulatta*, *M. nemestrina* and *M. leonina* were served as outgroups separated from the *M. fascicularis* group.

Subspecies identification and hybrid classification by SRY phylogenetic analysis

Combining the results of both HVSI and SRY phylogenetic analyses enabled comprehensive subspecies and hybrid classification of all 30 samples (**Table 3**). As for subspecies, all *Mfa* (TPK) samples were consistently assigned to the *Mfa* lineage by both maternal and paternal markers. All *Mff* samples were grouped within the *Mff* clade in the HVSI tree; however, in the SRY tree, PPT151, PPT165, PPT185, PPT195, and PPTC216 clustered within the pink clade along with the *Mff* reference, confirming *Mff* paternal ancestry. As for hybrid classification, the HVSI tree placed all KRI samples within a distinct intermediate clade separate from both parental subspecies. In the SRY tree, hybrid male samples showed differential paternal ancestry: KRI014 clustered within the *Mfa* (blue), suggesting *Mfa* paternal origin, while KRI005,

KRI006, KRI012, and KRI023 clustered within the pink clade alongside *Mff* samples, indicating *Mff* paternal ancestry. This divergence in Y-chromosomal lineage among hybrid individuals reflects the complexity of hybridization and male-mediated introgression in this population. ONT-based subspecies identification was concordant with morphological classification for all samples, with every individual assigned to the same subspecies group by both approaches, thereby confirming the reliability of the genetic workflow.

Discussion

This study demonstrates that faecal swab samples yield sufficient DNA quality and quantity for reliable subspecies identification of *Macaca fascicularis*. Mean DNA concentrations of 37.6 ± 17.7 ng/ μ L and A260/A280 ratios within the acceptable range (1.6–2.2) indicate effective DNA preservation under field conditions. NAPSeq buffer preserves nucleic acids at ambient temperatures for up to 10 days,⁽¹⁷⁾ addressing a critical limitation in remote field settings where cold chain infrastructure is unavailable. Successful amplification of both HVSI and SRY amplicons from all samples further confirms the suitability of faecal swab-derived DNA for genetic analysis. Because faecal samples deposited on the ground can be collected opportunistically without animal capture, this approach minimizes animal stress and reduces the risk of zoonotic pathogen transmission to researchers,^(2, 3) supporting non-invasive sample collection for subspecies identification of macaque population in ecological studies.

Phylogenetic analysis of HVSI sequences determined three distinct maternal clusters corresponding to *Mff*, *Mfa*, and hybrid individuals, consistent with previously reported maternal lineages in Thailand.⁽⁴⁾ The clustering of hybrid samples into a distinct clade is consistent with maternal inheritance of a hybrid-derived mtDNA haplotype, as the hybrid reference sequences used in this study were derived from individuals previously documented as *Mff* $\times$ *Mfa* hybrids within the same geographic contact zone.⁽¹²⁾ The overall mean percent identity of HVSI consensus sequences was $94.06 \pm 3.68\%$, reflecting expected sequence divergence between study samples and their corresponding reference sequences rather than sequencing error.^(5, 12) SRY-based analysis successfully discriminated *Mff* and *Mfa* paternal lineages with high sequence fidelity (mean percent identity of $99.94 \pm 0.08\%$). However, bootstrap support for the *Mff* clade

was relatively low, likely due to the conserved of Y-chromosomal sequences among closely related *Macaca* taxa.^(13, 14) Importantly, combining both markers revealed differential paternal ancestry among one sample hybrid males carried an *Mfa* paternal lineage, whereas other hybrid males carried *Mff* paternal lineages, reflecting asymmetric male-mediated gene transfer within the hybrid zone.^(5, 6) These findings demonstrated that the combined use of maternal and paternal markers is essential for comprehensive hybrid characterization, as neither marker alone can fully resolve ancestry.

The application of ONT provides several practical advantages over conventional Sanger or short-read sequencing platforms for field applications.^(11–13) The MinION device is compact, USB-powered, and capable of generating sequencing data in real time, enabling same-day subspecies identification without requirement of centralized laboratory infrastructure.^(15, 16) In addition, the ONT-based workflow described here supports multiplexed sequencing, allowing simultaneous processing of multiple samples within a single run. To our knowledge, this study represents the first application of portable nanopore sequencing for subspecies identification in *M. fascicularis*.

Complete concordance between ONT-based and morphological identification across all samples supports the validity of the genetic workflow for reliable subspecies classification. This agreement represents an essential step in method validation, demonstrating that the workflow accurately identifies individuals. Morphological assessment in this study was performed by experienced veterinarians through direct physical examination of anesthetized animals, enabling detailed evaluation of taxonomically relevant traits under controlled conditions. Nevertheless, morphological identification alone remains unreliable in hybrid zones, particularly for individuals with subtle or intermediate phenotypes. Notably, morphological assessment cannot resolve the differential paternal ancestry detected among hybrid males in this study, highlighting the additional resolution provided by molecular approaches. Therefore, molecular verification is recommended when selecting animals for biomedical research, where genetic background may influence experimental outcomes.^(5–7) In situations where animals cannot be approached or captured, a non-invasive genetic workflow described here provides reliable subspecies identification beyond the capability of morphological assessment.

This study has several limitations. The relatively small sample size ($n = 30$) from three geographic locations may not fully represent the genetic diversity across the whole country. Therefore, future studies should be incorporated larger and geographically broader sample sets to evaluate the performance of ONT-based genetic workflow using faecal samples collected directly from the ground to support field deployment.

Conclusion

This study establishes a practical, field-deployable genetic workflow for non-invasive subspecies identification of *M. fascicularis* using fecal swab samples and Oxford Nanopore Technologies. The ONT-based workflow offers key advantages over conventional sequencing methods, including portability, real-time data generation, multiplexed sample processing, reduced cost and turnaround time, making it particularly well-suited for ecological surveillance and quality control of non-human primate models in biomedical research.

Acknowledgements

The authors did not receive any financial support, funding, or grant from public, commercial, or not-for-profit organizations for the conduct of this study.

Conflicts of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data sharing statement

Statistical summaries are presented in the published article. The data sets generated or analyzed during the present study are available from the corresponding author on reasonable request.

References

1. Fan Z, Zhou A, Osada N, Yu J, Jiang J, Li P, et al. Ancient hybridization and admixture in macaques (genus *Macaca*) inferred from whole genome sequences. *Mol Phylogenet Evol* 2018; 127:376–86.
2. Khanal L, Chalise MK, Fan PF, Kyes RC, Jiang XL. Multilocus phylogeny suggests a distinct species status for the Nepal population of Assam macaques (*Macaca assamensis*): implications for evolution and conservation. *Zool Res* 2021; 42:3–13.
3. Kanthaswamy S, Satkoski J, George D, Kou A, Erickson BJ, Smith DG. Interspecies hybridization and the stratification of nuclear genetic variation of rhesus (*Macaca mulatta*) and long-tailed macaques (*Macaca fascicularis*). *Int J Primatol* 2008; 29:1295–311.
4. Malaivijitnond S, Hamada Y. Current situation and status of long-tailed macaques (*Macaca fascicularis*) in Thailand. *Trop Nat Hist* 2008; 8:185–204.
5. Gardner MB, Luciw PA. Macaque models of human infectious disease. *ILAR J* 2008; 49:220–55.
6. Ebeling M, Küng E, See A, Broger C, Steiner G, Berrera M, et al. Genome-based analysis of the nonhuman primate *Macaca fascicularis* as a model for drug safety assessment. *Genome Res* 2011; 21:1746–56.
7. Uno Y, Uehara S, Yamazaki H. Genetic polymorphisms of drug-metabolizing cytochrome P450 enzymes in cynomolgus and rhesus monkeys and common marmosets in preclinical studies for humans. *Biochem Pharmacol* 2018; 153:184–95.
8. Fooden J. Provisional classifications and key to living species of macaques (primates: *Macaca*). *Folia Primatol (Basel)* 1976; 25:225–36.
9. Fooden J. Taxonomy and evolution of the Sinica group of macaques: I. Species and subspecies accounts of *Macaca sinica*. *Primates* 1979; 20:109–40.
10. Hamada Y, Suryobroto B, Goto S, Malaivijitnond S. Morphological and body color variation in Thai *Macaca fascicularis fascicularis* north and south of the Isthmus of Kra. *Int J Primatol* 2008; 29:1271–94.
11. Phadphong P, Kanthaswamy S, Oldt RF, Hamada Y, Malaivijitnond S. Population structure of *Macaca fascicularis aurea*, and their genetic relationships with *M. f. fascicularis* and *M. mulatta* determined by 868 RADseq-derived autosomal SNPs: a consideration for biomedical research. *J Med Primatol* 2022; 51:33–44.
12. Bunlungsup S, Imai H, Hamada Y, Gumert MD, San AM, Malaivijitnond S. Morphological characteristics and genetic diversity of Burmese long-tailed macaques (*Macaca fascicularis aurea*). *Am J Primatol* 2016; 78:441–55.
13. Bunlungsup S, Imai H, Hamada Y, Matsudaira K, Malaivijitnond S. Mitochondrial DNA and two Y-chromosome genes of common long-tailed macaques (*Macaca fascicularis fascicularis*) throughout Thailand and vicinity. *Am J Primatol* 2017; 79:1–13.
14. Rovie-Ryan JJ, Khan FAA, Abdullah MT. Evolutionary pattern of *Macaca fascicularis* in Southeast Asia inferred using Y-chromosomal gene. *BMC Ecol Evol* 2021; 21:26.
15. Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol* 2021; 39:1348–65.
16. Patterson EC, Morrison-Lanjouw S, Jobling MA, Wetton JH. Rapid on-site universal vertebrate species identification via multi-barcode nanopore sequencing. *PLoS One* 2025; 20:e0336383.
17. Menke S, Gillingham MAF, Wilhelm K, Sommer S. Home-made cost effective preservation buffer is a better alternative to commercial preservation methods for microbiome research. *Front Microbiol* 2017; 8:102.