Genetic Assessment of the Sunda Pangolin (*Manis javanica*) Using Degraded and Fresh Samples

Project Report

Đỗ Văn Thao[1], Nguyễn Văn Thành[1], Dương Thúy Hà[1], và Lê Đức Minh[1,2]

[1]Hanoi University of Science
Vietnam National University
334 Nguyễn Trai Road, Thanh Xuan District
Hanoi, Vietnam

[2]Centre for Natural Resources and Environmental Studies
Vietnam National University
19 Le Thanh Tong Street, Hoan Kiem District
Hanoi, Vietnam
Email: minh.le.cres@gmail.com

Project description

Pangolins of the genus *Manis* have been heavily exploited for food and traditional medicine in many countries in Southeast Asia. As a result, many populations of this group have declined drastically over the last two decades to the point that it is difficult to survey the species in the wild using conventional methods. In addition, many trade products originating from pangolins are scales, which are difficult to identify to species level for trade monitoring and law enforcement. It is therefore important to determine good laboratory protocols to effectively extract DNA from degraded samples, which can be collected from the field and/or confiscated from the trade.

To this end, we extracted DNA from scale, dung, hair, and blood, which were mostly collected from individuals kept at Cuc Phuong National Park. We tried different available protocols, and optimized them for the pangolin samples. After sequencing the successfully extracted samples, we analyzed the obtained data along with data available sequences from GenBank using phylogenetic and population aggregation approaches to examine genetic diversity of this group.

Material and methods

Molecular data and laboratory protocols

Twenty two newly collected samples from ten individuals were used in this study. The details of the samples were shown in Table 1. Samples originated from individuals kept in the Carnivore and Pangolin Conservation Program (CPCP) at Cuc Phuong National Park, except one collected from Bu Gia Map National Park.

Table 1. Details of the samples used in this study

Number	Sample codes	Sample type	Source	Year collected
1	F1a	Scale	CPCP	2012
2	F1b	Haif	CPCP	2012
3	F1f	Dung	CPCP	2012
4	F2a	Scale	CPCP	2012
5	F2b	Hair	CPCP	2012
6	F2f	Dung	CPCP	2012
7	F3a	Scale	CPCP	2012

8	F3b	Hair	CPCP	2012
9	F4a	Scale	CPCP	2012
10	F4b	Hair	CPCP	2012
11	F4f	Dung	CPCP	2012
12	F5a	Scale	CPCP	2012
13	F5b	Hair	CPCP	2012
14	F5f	Dung	CPCP	2012
15	F6a	Scale	CPCP	2012
16	F6b	Hair	CPCP	2012
17	F6f	Dung	CPCP	2012
18	F7a	Scale	CPCP	2012
19	F7b	Hair	CPCP	2012
20	F8f	Dung	CPCP	2012
21	F9f	Dung	CPCP	2012
22	F10m	Tissue from ear	Bu Gia Map National Park	2010

We sequenced the partial control region for all 22 samples using two primers published by Hsieh et al. (2011). Additional 54 published sequences of *Manis* sp. and *Manis pentadactyla* were included in the analyses. *Manis pentadactyla* was used to provide outgroup polarity.

DNA was extracted using DNeasy blood and tissue kit (Qiagen, California) following the manufacturer's instruction. Extracted DNA was amplified by PCR mastermix (Fermentas, Canada). The PCR volume consisted of 21µl (10µl of mastermix, 5µl of water, 2µl of each primer at 10pmol/µl and 2µl of DNA or higher depending on the quantity of DNA in the final extraction solution). PCR condition was: 95°C for 5 minutes to activate HotStarTaq; with 40 cycles at 95°C for 30s, 45° for 45s, 72°C for 60s; and the final extension at 72°C for 6 minutes. Negative controls were used in all amplifications to check for possible contamination. Successful amplifications were purified to eliminate PCR components using GeneJET™ PCR Purification kit (Fermentas, Canada). Purified PCR products were sent to Macrogen Inc. (Seoul, South Korea) for sequencing.

For hair and scale, samples were cleaned with Clorox 10% for 5 minutes to avoid potential contamination on the surface. The samples were then washed with purified water

several times, and put on clean surface until they were dry. The cleaned samples were cut into small pieces to facilitate the lysis process, before being extracted as fresh tissue and blood samples.

For dung, samples were centrifuge at 6000rpm for 5 minutes to remove fibers. The samples were then centrifuged at 6000rpm for 5 minutes. The supernatant was then removed from the samples. The samples were dried up in a refrigerator at 4^{o}C overnight to take out remaining alcohol. ASL buffer was added to the samples and then vortex for 5 minutes. The samples were then incubated in a refrigerator for four days. During this period, samples were vortex occasionally. The samples were then centrifuged at 6000rpm speed for 5 minutes. After deposit was removed, a InhibitEx tablet was added to the supernatant. The mixture was then vortex for three minutes to dissolve. It was then incubated at room temperature for 5 minutes, and centrifuged at 6000rmp for 6 minutes. The supernatant was then used for extraction as for fresh tissue or blood samples.

Phylogenetic analyses

The sequences were aligned in BioEdit v7.1.3 (Hall 1999) with default settings. Data were analyzed using maximum parsimony (MP) as implemented in PAUP 4.0b10 (Swofford 2001) and Bayesian analysis as implemented in MrBayes 3.2.1 (Huelsenbeck and Ronquist 2001). Uncorrected pairwise divergences were calculated in PAUP 4.0b10. (supplementary data). All settings were followed Le et al. (2006), except that the number of generations in the Bayesian analyses was increased to 10^{7}.

Results and discussion

Twenty two samples of the Sunda Pangolin were successfully sequenced. The final matrix consisted of 38 ingroup and 18 outgroup terminals with 580 aligned characters. In the single-model Bayesian analysis, ln*L* scores reached equilibrium after 9,000 generations in both runs.

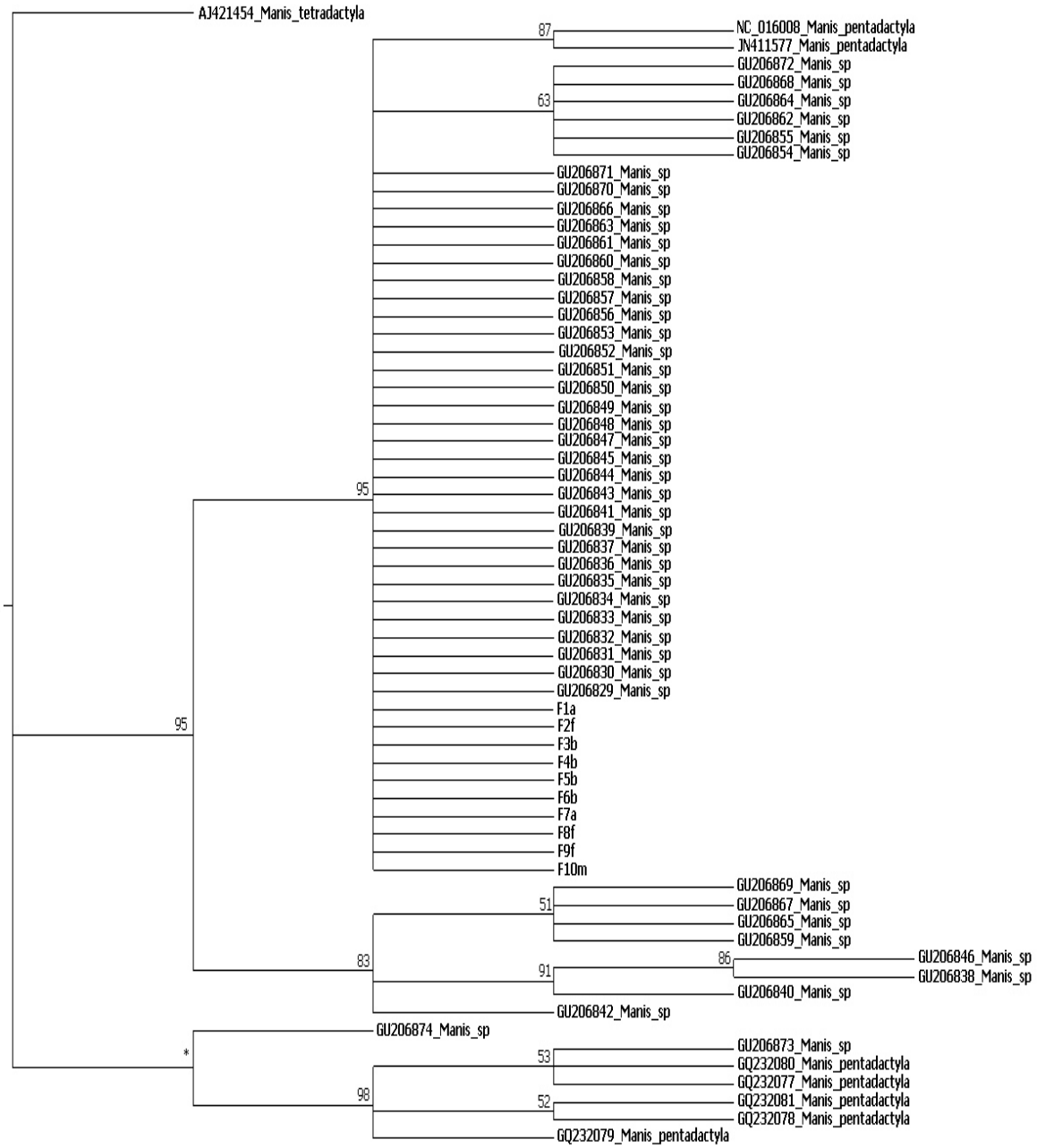

Figure 1. Strict consensus parsimony tree based on 580 alligned characters of control region (Tree length=159, Consistency index=0.86, Retention index=0.94). Values above branches are boostrap based on 1000 replicates.

Tree topologies obtained from MP and combined Bayesian analyses are almost identical in resolving important nodes (Fig. 1 and Fig. 2). Overall, samples from were grouped with most samples identified as *Manis* sp. and two samples identified as *M. pentadactyla* with strong support from both analyses. These samples are virtually identical in terms of sequence data (Table 2). From the evidence, we can conclude that they belong to the same species. Since our samples came from specimens with clear morphological identification, we can confirm the these samples belong to the Sunda Pangolin (*Manis javanica*).

Figure 2. Bayesian consensus tree based on 580 aligned characters of control region. General Time Reversible (GTR) was selected as optimum model for the Bayesian runs.

Table 2. Site difference in nucleotide sequences of *Manis javanica* used in this study

Haplotype[b]	N[c]	Nucleotide site[a]
		1111111111111111111222222222233333333344444445555 1112244566688891233444455566677789124445567011788992248890013 4560345612156847843108058902568912632983462617258035023485742306
GU206829	13	- ACT- ACTCCTTT -GGGAGTTTTCGTA - AGTAATATGTTAGTACGGCTCTCGTGCGATAGCTTA
GU206831	4	. . . . . G . . . . . T . . . . .
GU206834	1	. . . . . G . . . . . T . . . . .
GU206835	1	. . . . . T . . . . . T . . . . .
GU206837	3	. . . . . T . . . . .
GU206838	2	. . . . . C . C . . A T CC C . G AG . . C . . . . . A C . . A . . . . . T . . . G . . . . .
GU206840	1	. . . . . C . . A T CC C . G . G . . C . . . . . A C . . A . . . . . T . . . G . . . . .
GU206841	2	. . . . . T . . . . .
GU206842	1	. . . . . A T CC T . G . G . . . . . A C . . . . . G . . . . .
GU206847	1	. . . . . A . . G . . T . . . C . . . . . G . . . C . . . . . T . . . . .
GU206848	1	. G . . . . . G . . . . . T . . . . .
GU206851	1	. . . . . T . . . . .
GU206854	6	. . . . . A . . . . .
GU206857	1	. . . . . T . . . . . G . . . . . C . . . . .
GU206859	1	. . . . G . . . . . C . . A T CC T . G . G . . . . . A C . T . . . . . T . . . G . . . . .
GU206860	1	. . . . . G . . G . . . . .
GU206865	1	. . . . G . . . . . C . A T C C T . G . G . . . . . A C . T . . . . . T . . . G . . . . .
GU206867	1	. . . . G . . . . . C . . A T C C T . G . G . . . . . A C . . . . . T . . . G . . . . .
GU206869	1	. . . . G . . . T . . . . . CA T C C T . G . G . . . . . A C . T . . . . . T . . . G . . . . .
GU206871	1	. . . . . T . . . . .
GU206873	1	T . T . T . AA . T . C CA . A . C . . . . A T . C C CG . . . . . AC . . . T . . C TC T GA TC A T ATC . A T ACC
GU206874	1	. . T - T . AA . T . C C- . . . C . . . . A T . C C - G . . . . . C . GC . A . T . . C . C T GA TC A T AT . . A T ACC
GQ232077	1	T . T . T . AA . T . C CA . A . C . . . . A T . C C CG . . . . . AC . . . T T . C TC T GA TC A T ATC . A T ACC
GQ232078	1	T . T . T . AA . T C C CA . A . C . . . . A T . C CCG . . . . . AC . A . T T . C TC T GA TC A T ATC . A T ACC
GQ232079	1	T . T . T . AA . T . C CA . A . C . . . . A T . C CCG . . . . . AC . A . T . . C TC T GA TC A T ATC . A T A CC

GQ232080	1	T . T . T . AA . T . C CA . ACC . . . . A T . C CCG . . . . . .AC . . . T T . C TC T GA TC A T ATC . A T ACC
GQ232081	1	T . T . T . AA . T C C CA . A .C . . . . A T . C CCG . . . . . G . . AC . A . T . . . TC T GA TC A T ATC . A T A CC
F1	1	. . . . .
F2	1	. . . . . .G. . . . . T. . . . .
F3	1	. . . . .
F4	1	. . . . .
F5	1	. . . . . .C . . . . .
F6	1	. . . . .
F7	1	. . . . .
F8	1	. . . . .
F9	1	. . . . .
F10	1	. . . . .

This study for the first time provides assessment of genetic study of the Sunda pangolin populations in Vietnam. Although samples often do not have accurate localities, comparing genetic data from these samples and sequences available on GenBank, which originated from Taiwanese wildlife markets, show that this species has a low level of genetic diversity. We suggest that a range wide genetic study of this species be conducted to further explore the issue. Finally, the laboratory protocols successfully used in this project can be applied to future surveys where non-invasive samples can be collected in the field.

Acknowledgements

This study was supported by the National Foundation for Science and Technology Development of Vietnam (NAFOSTED: Grant Nos. 106.15-2010.30 to M. Le). The Carnivore and Pangolin Conservation Program at Cuc Phuong National Park generously provided funding for research activities.

References

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium 41: 95-98.

Hsieh, H.-M., Lee, J. C.-I., Wu, J.-H., Chen, C.-A., Chen, Y.-J., Wang, G.-B., Chin, S.-C., Wang, L.-C., Linacre, A., Tsai, L.-C. (2011). Establishing the pangolin mitochondrial D-loop sequences from the confiscated scales. Forensic Science International 20: 303–307.

Huelsenbeck, J. P. and Ronquist, F. 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics 17: 754-755.

Le, M., Raxworthy, C. J., McCord, W. P., and Mertz, L. 2006. A molecular phylogeny of tortoises (Testudines: Testudinidae) based on mitochondrial and nuclear genes. Molecular Phylogenetics and Evolution 40: 517-531.

Swofford, D.L. 2001. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods), version 4. Sinauer Associates, Massachusetts.