Genomic insights into conservation status of *Manis pentadactyla pentadactyla* populations in Taiwan

Ting-Hsuan Kuo[1], Pei-Wei Sun[2], Chien-Yu Chen[3], Si-Wei Cheng[4], Yu-Ten Ju[5]

1 國立台灣大學動物科學所碩士班研究生
2 國立師範大學生命科學所碩士班研究生
3 國立台灣大學生物機電工程學系教授
4 行政院農委會特有生物研究保育中心研究員
5 國立台灣大學動物科學系教授

Introduction

Pangolins (monotypic order Pholidota and family Manidae) are nocturnal and scale-covered mammals that prey on ants and termites. Pangolins act as a control agent for local ant and termite abundance and community structure, thus affecting ecosystem function (Challender et al., 2019). There are eight species of pangolin, four of which are distributed within Africa (Tree pangolin *Phataginus tricuspis*, Long-tailed pangolin *P. tetradactyla*, Temminck's ground pangolin *Smutsia temminckii* and Giant pangolin *S. gigantean*) and four over Asia (Malayan pangolin *Manis javanica*, Chinese pangolin *M. pentadactyla*, Indian pangolin *M. crassicaudata* and Philippine pangolin *M. culionensis*) (Gaubert et al., 2018; Gaudin et al., 2009) have been listed in 'Appendix I' of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) (Heinrich et al., 2016). They are considered the most trafficked mammals in the world which accounts for about 20% of all transnational wildlife trade (Challender et al., 2016; Cheng et al., 2017; Nijman et al., 2016; Wu et al., 2007). Overexploitation driven by wildlife trade for their meat as delicacy or scales for traditional medicine which has no proven pharmacological effects almost pushed them to the edge of extinction (Challender, 2011; Hsieh et al., 2011; Liu and Weng, 2014; Luczon et al., 2016; Zhou et al., 2014). Although pangolins are protected by national and regional laws, wildlife poaching and international trades have been still reported and population declines continues (Heinrich et al., 2016; Cheng et al., 2017; Challender, 2011; Zhou et al., 2014; Hu et al., 2020).

Manis pentadactyla pentadactyla is a subspecies of Chinese pangolin found only on the sub-tropical island Taiwan (Schlitter et al., 1993). They live primarily in agricultural fields on mountain slopes below 1,000m a.s.l., with the highest densities of individuals being at about 300m (Lin, 1982). They were widely distributed throughout lowland Taiwan in the late nineteenth century and mid-twentieth century (Swinhoe, 1870; Horikawa, 1932; Chao, 1989). However, it has been estimated that as many as 60,000 individuals were harvested annually from 1950 to 1970 for the local traditional medicine and game meat markets (Chao, 1989). Recently, the study not only revealed their mating system but low genetic diversity inferred by microsatellite markers.

Bottleneck analysis suggests that the Chinese pangolins have suffered a reduction in population size, subsequently causing a decline in genetic diversity (Sun et al., 2020). The over-exploitation and demands of the leather industry in Taiwan between 1950 and 1980 are likely major factors that impacted the genetic bottleneck (Chao, 1989). Hu et al. (2020) also found that the genetic variation of the Taiwanese population is significantly lower than that of mainland China, suggesting that the Taiwanese population is a small founding population migrated from mainland China during the late Pleistocene (Fig. 1a). Moreover, Taiwan pangolin population went through a steep population decline (Fig. 1b) and would face with inbreeding crisis (Fig. 3) (Hu et al., 2020). Taiwan pangolin population might shrinked not because of anthropogenic influence which caused bottleneck effect but small population migration to Taiwan ended up with founder effect. Therefore, conservation management of small population for Taiwan pangolin is more significant than before.

Small and fragmented populations is vulnerable because of environmental effects (e.g. disease and habitat destruction), demographic stochasticity and detrimental genetic effect. The role of genetic factor in the long-term viability of small populations has attracted much attention over past few decades (Caughly, 1994). An increasing number of empirical studies on critically endangered species indicates that small populations are often exposed to genomic erosion (van der Valk et al., 2019; Feng et al., 2019), which increase inbreeding level and reduce species viability through gene drift (Keller, 2002; Lande and Shannon, 1996; Bijlsma and Loeschcke, 2012). The influence of these process can vary among population depending on their different

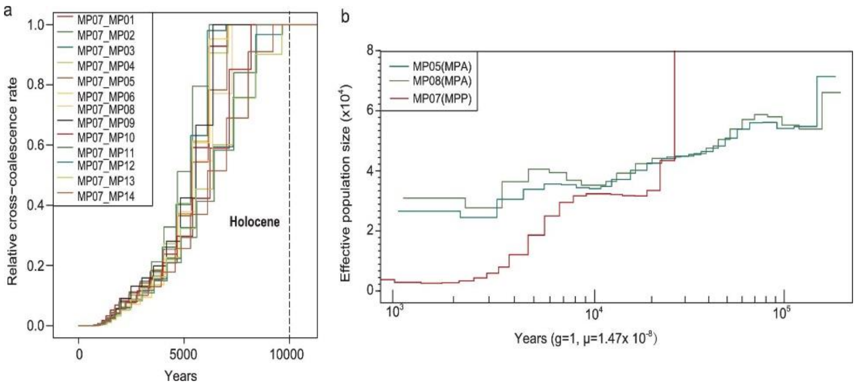

Figure1. Demographic history of the Taiwan pangolin. (a) Divergence time of the Taiwan pangolin (MP07) from its ancestral MPA group (MP01–MP14). (b) Population size dynamics of the Taiwan pangolin (MP07) and non-Taiwan MPA. Two non-Taiwan MPA individuals (MP05 and MP08) that have the highest

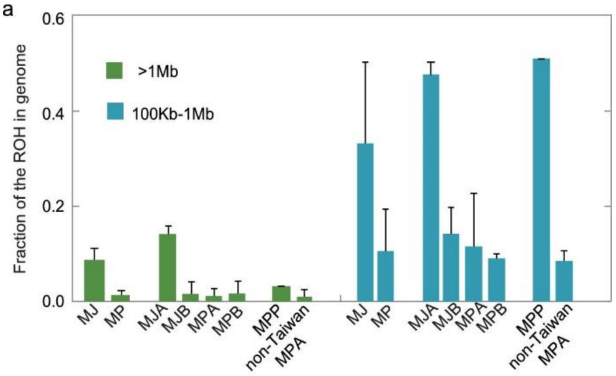

Figure 2. Inbreeding levels and genetic load of pangolins. Error bars represent SD. (a) Fraction of the runs of homozygosity (ROH) in the genome, including the medium (100 kb–1 Mb) and long (>1 Mb) ROH, of MJ, MP, MJA, MJB, MPA, MPB, MPP and non-Taiwan MPA.

demographic histories (e.g., population fluctuations and founder effects), sensitivity to environmental changes or life-history traits. In addition, population fitness will also relate to rare large-effect deleterious alleles and numerous small-effect deleterious alleles among the founding individuals of these populations. Therefore, predicting the degree of genomic erosion and quantifying level of inbreeding that threatened populations are exposed to is thus challenging, but it indicates more implication for conservation developing management strategies (Díez-del-Molino et al., 2018). Genetic data can help define management units (Palsboll et al., 2007) and identify the individuals most likely to lead to genetic rescue in translocation programmes (Whitely et al., 2015). Enhancing gene flow is considered a powerful conservation tool for reducing genomic erosion in a range of threatened species, because it can reduce the expression of recessive or partially-recessive deleterious alleles in hybrids of the receiving population (Whitely et al., 2015; Frankham, 2015).

The genomic revolution has democratized the field of population genomics, allowing high-throughput sequencing to be applied in nearly any organism, including natural populations of rare wildlife species. Population genomics in wildlife can inform management and conservation efforts, such as delineating population units for conservation, maintaining genetic diversity or predicting adaptive potential. The importance of genetic variation in conservation, including its role in setting conservation targets and monitoring the status of biodiversity, is increasingly recognized and can benefit from genomics tools (Hohenlohe and Rajora, 2020). With fine-scale genomic data, mapped to a reference genome, it is possible to identify runs of homozygosity (ROH) – chromosomal regions that have few or no heterozygous nucleotide sites because both chromosome copies derive from a single copy in a relatively recent common ancestor (Ceballos et al., 2018). The proportion of the genome that is in ROH, or identical by descent, has long been central to the concept of inbreeding, because it is the result of relatedness between parents. Being able to map these regions in the genome reveals several novel insights that illustrate the power of population genomics approaches. In addition, ROH provide precise estimates of individual-level inbreeding which are more accurate than other methods (Kardos et al., 2015). The relative locations of ROH among individuals and populations can be informative for controlled breeding or genetic rescue attempts. For example, if two individuals share ROH at the same chromosomal region due to common ancestry, their offspring will also have those regions of reduced diversity. However, if two individuals have different ROH, mating between them can produce offspring with lower inbreeding coefficients, potentially relieving inbreeding depression. By statistically powerful genomic data, comparison of inbreeding level is essential for developing management strategies.

Earlier studies focused on Taiwan pangolin population structure by using mitochondrial genes (cytochrome b and control region), major histocompatibility complex (MHC) or microsatellite. Wang et al. (2007) obtained cytochrome b and control region sequence from 66 samples which identified 22 haplotype and indicated northern, central, southern and eastern Taiwan populations having their own unique haplotypes. Liu et al. (2017) amplified MHC sequences of 92 individuals from central and northern Taiwan, however, they found MHC sequence wasn't suitable to study population genetics of Taiwan pangolin because of its haplotype diversity was low. Chang et al. (2014) combined 10 microsatellite markers with radio tracking to reveal population structure in Taitung Luanshan area. Wu et al. (2015) added microsatellite markers to 15 and indicated that northern Taiwan pangolin population had low inbreeding level. Based on the same 15 microsatellite markers, Liu et al. (2017) found that there were difference of population structure in central-southern and northern Taiwan pangolin population. However, not only their population structure results were inconsistent but inbreeding level and recessive alleles of Taiwan pangolin population remained unknown.

Specific aims

1. Reveal population structure of *Manis pentadactyla* from Taiwan and China

Thirty-one *Manis pentadactyla* from Taiwan raw pair-end read data sequence from Illumina NovaSeq 6000 platform. Twenty-one *Manis pentadactyla* from China raw pair-end read data will download from NCBI by SRAtools kits. For quality trimming and adaptor clipping, we used Trimmomatic 0.38 (Bolger et al., 2014) for Illumina paired-end reads removing adapters, leading low quality or N bases below quality 3 and trailing low quality or N bases below quality 3; Scanned the read with a 4-base wide sliding window and cut when the average quality per base drops below 15.

Dropped reads below the 100 bases long. We checked the trimmed reads quality by FASTQC (Andrews, 2010) report.

The trimmed paired-end reads were aligned to an assembled Chinese pangolin (GenBank assembly accession: GCA_014570555.1) by BWA 0.7.17 (Li, 2013). Binary sequence alignment files were generated by SAMtools v.1.3 (Li et al., 2009). Low mapping quality (MAPQ <10) and PCR duplicate were excluded using Sambamba 0.7.0 (Artem et al., 2015). The filtered sam files were conducted SNP-calling analysis by GATK 4.1.0.0 (Van der Auwera et al., 2013).
(In this part, we require about 10TB storage.)

2. Compare inbreeding level between populations

To estimate heterozygosity, VCF merged file translate VCF files into arlequin project (*.arp) files by PGDspider (Lischer and Excoffier, 2012). Arlequin project file analyzed mean heterozygosity by Arlequin 3.5.2.2 (Excoffier and Lischer, 2010).

To estimate ROH, merged VCF file of populations (Danecek et al., 2011) and processed quality control of SNPs by VCFtools (Danecek et al., 2011) first. The criteria of ROH selection using PLINK1.9 (Chang et al., 2015). F_{roh} was defined as individual total length of ROH divided by total length of genome. The visualization was plotted by R (Ihaka and Gentleman, 1996).
(In this part, we required 500GB for calculation.)

3. Loss of function genes, gene enrichment and protein structure in populations

The vcf files which contained SNPs found from ROH were obtained the same genotype in population by VCFtools (Danecek et al., 2011) and added loss of function tags by Snpeff (Cingolani et al., 2012). Filtered out genes loss of function effect less than 50% transcript using SnpSift (Ruden et al., 2012).

The gene lists which annotated from high impact variant predicted by Snpeff (Cingolani et al., 2012) were input to KEGG Orthology Based Annotation System (KOBAS) (Bu et al., 2021) for gene enrichment analysis.For observing protein structure difference of wild type and LoF, they were predicted by Alphafold (Jumper et al., 2021), which demonstrate a new architecture to jointly embed multiple sequence alignments (MSAs) and pairwise features.

Broader impact

The recessive allele in specific population might cause disease or influence their viability, thus, the finding would be critical for conservation management for *Manis pentadactyla pentadactyla*.

Reference

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Bijlsma, R., & Loeschcke, V. (2012). Genetic erosion impedes adaptive responses to stressful environments. Evolutionary Applications, 5(2), 117-129.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114-2120.

Bu, D., Luo, H., Huo, P., Wang, Z., Zhang, S., He, Z., ... & Kong, L. (2021). KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. Nucleic acids research, 49(W1), W317-W325.

Cai, B. Q. (1999). The later Pleistocene fossil mammals and the palaeogeographical environment of the Taiwan Strait. J Xiamen Univ, 4, 29-33. Caughley, G. Directions in conservation biology. J. Anim. Ecol. 63, 215 (1994).

Ceballos, G., Ehrlich, P. R., & Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. Proceedings of the national academy of sciences, 114(30), E6089-E6096.

Challender, D. W., & Hywood, L. (2011). Asian pangolins: increasing affluence driving hunting pressure. Traffic Bulletin, 23(3), 92-93.

Challender, D. W., Harrop, S. R., & MacMillan, D. C. (2015). Understanding markets to conserve trade-threatened species in CITES. Biological Conservation, 187, 249-259.

Challender, D. W., Nash, H., & Waterman, C. (2019). Pangolins: science, society and conservation. Academic Press.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience, 4(1), s13742-015.

Chang, S.P., 2014. The kinship and social structure of the Formosan pangolin (Manis pentadactyla pentadactyla) in Luanshan, Taitung, based on microsatellite variations. M.Sc. Thesis. National Pingtung University of Science and Technology.

Chao, J. T. (1989). Studies on the conservation of the Taiwanese Pangolin (Manis pentadactyla pentadactyla). General biology and current status.

Cheng, W., Xing, S., & Bonebrake, T. C. (2017). Recent pangolin seizures in China reveal priority areas for intervention. Conservation Letters, 10(6), 757-764.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., ... & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly, 6(2), 80-92.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. Bioinformatics, 27(15), 2156-2158.

Díez-del-Molino, D., Sánchez-Barreiro, F., Barnes, I., Gilbert, M. T. P., & Dalén, L. (2018). Quantifying temporal genomic erosion in endangered species. Trends in Ecology & Evolution, 33(3), 176-185.

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular ecology resources, 10(3), 564-567.

Feng, S., Fang, Q. I., Barnett, R., Li, C., Han, S., Kuhlwilm, M., ... & Zhang, G. (2019). The genomic footprints of the fall and recovery of the crested ibis. Current Biology, 29(2), 340-349.

Frankham, R. (2015). Genetic rescue of small inbred populations: Meta-analysis reveals large and consistent benefits of gene flow. Molecular ecology, 24(11), 2610-2618.

Gaubert, P., Antunes, A., Meng, H., Miao, L., Peigné, S., Justy, F., ... & Luo, S. J. (2018). The complete phylogeny of pangolins: scaling up resources for the molecular tracing of the most trafficked mammals on earth. Journal of Heredity, 109(4), 347-359.

Gaudin, T. J., Emry, R. J., & Wible, J. R. (2009). The phylogeny of living and extinct pangolins (Mammalia, Pholidota) and associated taxa: a morphology based analysis. Journal of mammalian evolution, 16(4), 235-305.

Heinrich, S., Wittmann, T. A., Prowse, T. A., Ross, J. V., Delean, S., Shepherd, C. R., & Cassey, P. (2016). Where did all the pangolins go? International CITES trade in pangolin species. Global Ecology and Conservation, 8, 241-253.

Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. Molecular Ecology, 30(1), 62-82.

Horikawa, Y. (1932). Field guide to the mammals in Taiwan. Society of Taiwan Museum Press, Taipei.

Hsieh, H. M., Lee, J. C. I., Wu, J. H., Chen, C. A., Chen, Y. J., Wang, G. B., ... & Tsai, L. C. (2011). Establishing the pangolin mitochondrial D-loop sequences from the confiscated scales. Forensic Science International: Genetics, 5(4), 303-307.

Hu, J. Y., Hao, Z. Q., Frantz, L., Wu, S. F., Chen, W., Jiang, Y. F., ... & Yu, L. (2020). Genomic consequences of population decline in critically endangered pangolins and their demographic histories. National science review, 7(4), 798-814.

Ian, C. W. (1979). An evaluation of the changes of bird and mammal populations after 6 years of hunting ban enforcement in Taiwan. Environmental Research Center. Taitung: Tung-Hai University Press.

Ihaka, R., & Gentleman, R. (1996). R: a language for data analysis and graphics. Journal of computational and graphical statistics, 5(3), 299-314.

Jost, L. (2006). Entropy and diversity. Oikos, 113(2), 363-375.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583-589.

Kardos, M., Qvarnström, A., & Ellegren, H. (2017). Inferring individual inbreeding and demographic history from segments of identity by descent in Ficedula flycatcher genome sequences. Genetics, 205(3), 1319-1334.

Kawamura, A., Chang, C. H., & Kawamura, Y. (2016). Middle Pleistocene to Holocene mammal faunas of the Ryukyu Islands and Taiwan: An updated review incorporating results of recent research. Quaternary International, 397, 117-135.

Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. Trends in ecology & evolution, 17(5), 230-241.

Lande, R., & Shannon, S. (1996). The role of genetic variation in adaptation and population persistence in a changing environment. Evolution, 434-437.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics, 27(21), 2987-2993.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... & Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics, 25(16), 2078-2079.

Lin LK. A study on the mammals in Taiwan. M.Sc. Thesis, Tung-Hai University. 1982.

Lin, J.S., 2010. Investigation of Intestinal Parasite in Formosan Pangolins (Manis pentadactyla pentadactyla) at Luanshan, Taitung. M.Sc. Thesis. National Pingtung University of Science and Technology.

Liu, J.T., 2017. Genetic diversity and genetic structure of Formosan pangolin in Taiwan based on MHC gene and microsatellite markers. M.Sc. Thesis. National Taiwan University.

Liu, Y., & Weng, Q. (2014). Fauna in decline: plight of the pangolin. Science, 345(6199), 884-884.

Luczon, A. U., Ong, P. S., Quilang, J. P., & Fontanilla, I. K. C. (2016). Determining species identity from confiscated pangolin remains using DNA barcoding. Mitochondrial Dna Part B, 1(1), 763-766.

McCullough DR. Status of larger mammals in Taiwan. Taipei: Tourism Bureau Press; 1974.

Nijman, V., Zhang, M. X., & Shepherd, C. R. (2016). Pangolin trade in the Mong La wildlife market and the role of Myanmar in the smuggling of pangolins into China. Global Ecology and Conservation, 5, 118-126.

Palsbøll, P. J., Berube, M., & Allendorf, F. W. (2007). Identification of management units using population genetic data. Trends in ecology & evolution, 22(1), 11-16.

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. Proceedings of the national academy of sciences, 105(4), 1118-1123.

Ruden, D. M., Cingolani, P., Patel, V. M., Coon, M., Nguyen, T., Land, S. J., & Lu, X. (2012). Using Drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, SnpSift. Frontiers in genetics, 3, 35.

Sun, N. C. M., Chang, S. P., Lin, J. S., Tseng, Y. W., Pei, K. J. C., & Hung, K. H. (2020). The genetic structure and mating system of a recovered Chinese pangolin population (Manis pentadactyla Linnaeus, 1758) as inferred by microsatellite markers. Global Ecology and Conservation, 23, e01195.

Swinhoe, R. (1870). Catalogue of the mammals of China (south of the Yangtsze) and of the island of Formosa. In Proc. Zool. Soc. London (Vol. 1870, pp. 615-653).

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J., & Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics, 31(12), 2032-2034.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., ... & DePristo, M. A. (2013). From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Current protocols in bioinformatics, 43(1), 11-10.

Van Der Valk, T., Díez-del-Molino, D., Marques-Bonet, T., Guschanski, K., & Dalén, L. (2019). Historical genomes reveal the genomic consequences of recent population decline in eastern gorillas. Current Biology, 29(1), 165-170.

Wang, P.J., 2007. Application of Wildlife Rescue System in Conservation of the Formosan Pangolin (Manis pentadactyla pentadactyla) M.Sc. Thesis. National Taiwan University.

Wang, Y. (1986). A study on the utilization of wild animal in Taiwan (I). Ecol Res Series, 11.

Whiteley, A. R., Fitzpatrick, S. W., Funk, W. C., & Tallmon, D. A. (2015). Genetic rescue to the rescue. Trends in ecology & evolution, 30(1), 42-49.

Wilson, D. E., & Reeder, D. M. (Eds.). (2005). Mammal species of the world: a taxonomic and geographic reference (Vol. 1). JHU press.

Wu, S. B., & Ma, G. Z. (2007). The status and conservation of pangolins in China. TRAFFIC East Asia Newsletter, 4, 1-5.

Wu, Y.H., 2015. Genetic Diversity and Genetic Structure of Formosan Pangolins in Taiwan Based on Microsatellite Markers and Mitochondrial DNA. M.Sc. Thesis. National Taiwan University.

Zhou, Z. M., Zhou, Y., Newman, C., & Macdonald, D. W. (2014). Scaling up pangolin protection in China. Frontiers in Ecology and the Environment, 12(2), 97-98.