**UNIVERSITI PUTRA MALAYSIA**

# *ISOLATION AND CHARACTERISATION OF MAJOR HISTOCOMPATIBILITY COMPLEX GENE CLASS I AND II IN MALAYAN TAPIR (Tapirus indicus Desmarest*

**NURUL ADILAH BINTI ISMAIL**

**FS 2020 35**

# ISOLATION AND CHARACTERISATION OF MAJOR HISTOCOMPATIBILITY COMPLEX GENE CLASS I AND II IN MALAYAN TAPIR (*Tapirus indicus* Desmarest)

By

**NURUL ADILAH BINTI ISMAIL**

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science**

**April 2019**

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

**ISOLATION AND CHARACTERISATION OF MAJOR HISTOCOMPATIBILITY COMPLEX GENE CLASS I AND II IN MALAYAN TAPIR (*Tapirus indicus* Desmarest)**

By

**NURUL ADILAH BINTI ISMAIL**

**April 2019**

**Chair : Geetha Annavi, PhD**
**Faculty : Science**

Malayan tapir is one of the five tapir species in the world and is listed as endangered in the IUCN Red List due to multiple threats such as habitat loss and human disturbance that lead to its high population decline. Low population number in this species increase the risk of inbreeding that could result in reduction in genome-wide genetic variation and particularly risky if it affects the gene responsible for immune response i.e. MHC gene. Class I and II MHC genes are responsible for encoding MHC molecules in the cells that recognise pathogenic peptides and present them to T-Cells on the cell surface for adaptive immune response. However, at present there is no study related to MHC gene in Malayan tapir yet. This study characterise the MHC peptide-binding region (PBR) of the MHC class I and II gene in Malayan tapir by isolating the DNA, amplify the targeted region by PCR, cloning and sequencing; investigate if there is evidence of balancing selection by calculating the rate of non-synonymous ($d_N$) and synonymous ($d_S$) substitutions using MEGA and PAML and study its relationship with homologous genes of other species based on phylogenetic tree construction using evolutionary model. In this study at least five α1 and four α2 of class I alleles, alongside two DRA, two DQA, three DRB and three DQB of class II alleles were isolated. The α1 and α2 domains of class I and DRβ domain of class II gene display evidence of selection with $d_N / d_S > 1$. A total of 24 codons within exon 2 DRB gene were found to be under selection with ten of the codons under positive selection sites (PSS) are part of the codons forming the Antigen Binding Site (ABS). Class I α1 and α2 sequences formed two separate groups on phylogenetic tree when compared to other species indicating possibility of two different loci. Within class II genes, all genes show species specific monophyletic group formation except for DRB genes with intersperse relationship in their phylogenetic trees which may indicate occurrence of trans-species polymorphism of allelic lineage. To maintain and improve the variation within the MHC gene, it is recommended to genotype tapir at the MHC gene to avoid mating with similar MHC allele and variation.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia
sebagai memenuhi keperluan untuk ijazah Master Sains

**PEMISAHAN DAN PENCIRIAN KOMPLEKS KEHISTOSERASIAN MAJOR GEN KELAS I DAN II PADA TAPIR MALAYA (*Tapirus indicus* Desmarest)**

Oleh

**NURUL ADILAH BINTI ISMAIL**

**April 2019**

**Pengerusi : Geetha Annavi, PhD**
**Fakulti : Sains**

Tapir Malaya adalah salah satu daripada lima spesis tapir di dunia dan disenaraikan sebagai haiwan terancam dalam Senarai Merah IUCN berikutan pelbagai ancaman seperti kehilangan habitat dan gangguan manusia yang mengakibatkan penurunan populasi yang tinggi. Jumlah populasi yang rendah dalam kalangan spesis ini meningkatkan lagi risiko pembiakbakaan dalam yang memungkinkan lagi pengurangan dalam variasi genetik sekitar genom dan terutamanya amat berisiko sekiranya ia mendatangkan kesan terhadap gen yang bertanggungjawab terhadap respons imun, contohnya gen KKM. Gen KKM Kelas I dan II bertanggungjawab dalam pengekodan molekul-molekul KKM di dalam sel yang bertindak mengenal pasti peptida berpatogen dan membawa ia ke Sel-Sel T pada permukaan sel untuk penyesuaian respons imun. Walau bagaimanapun, setakat ini tidak terdapat sebarang kajian yang berkaitan dengan gen KKM dalam kalangan tapir Malaya. Kajian ini mengelaskan sempadan peptida yang bergabung '*peptide-binding region (PBR)'* kepada gen KKM kelas I dan II dalam tapir Malaya dengan mengasingkan DNA, menggandakan sempadan yang disasarkan dengan PCR, pengklonan dan penjujukan; untuk menyiasat sekiranya wujud bukti akan pemilihan pengimbangan dengan mengira kadar pengganti ketidaksinoniman ($d_N$) dan kesinoniman ($d_S$) menggunakan MEGA dan PAML dan mengkaji hubungannya dengan gen berhomolog daripada spesis-spesis lain berdasarkan pokok/ranting filogenetik menggunakan model evolusi. Dalam kajian ini, terdapat sekurang-kurangnya lima α1 dan empat α2 alel kelas 1, bersama dua DRA, dua DQA, tiga DRB, dan tiga DQB alel kelas II yang telah diasingkan. Domain α1 dan α2 kelas I dan domain DRβ gen kelas II menunjukkan bukti pemilhan dengan $d_N / d_S > 1$. Sejumlah 24 kodon dalam ekson 2 gen DRB didapati berada di bawah pemilihan dengan 10 kodon di bawah tapak pemilihan positif (TPP) adalah sebahagian daripada kodon yang membentuk Tapak Ikatan Antigen (TIA). Turutan-turutan α1 dan α2 kelas I membentuk dua kumpulan berasingan pada pokok filogenetik apabila dibandingkan dengan spesis-spesis lain menandakan kemungkinan wujudnya dua loci berlainan. Manakala dalam gen

kelas II, kesemua gen menunjukkan pembentukan kumpulan monofiletik yang spesifik terhadap spesis kecuali gen DRB yang menunjukkan hubungan selerak dalam pokok filogenetiknya yang mungkin menandakan kewujudan polimorfisme spesis trans terhadap leluhur alel. Untuk mengekalkan dan menambah baik variasi dalam gen KKM, adalah dicadangkan untuk menggenotipkan tapir pada peringkat gen KKM bagi mengelakkan pengawanan dengan alel dan variasi KKM yang serupa.

**ACKNOWLEDGEMENTS**

All praise to Allah for giving me the opportunity and strength to complete this thesis. My utmost gratitude to my parents, family and friends for their unconditional supports and understanding. I'm grateful to Dr Geetha Annavi, for giving me the opportunity and guidance to complete this project. I'd also like to express my gratitude to my co-supervisors, Dr Simon, Dr Christina and Dr Wan Norhamidah for their supervision throughout the project. Special thanks to Ministry of Natural Resource and Environment for providing the permits, PERHLITAN, Sungai Dusun Wildlife reserve, and Zoo Negara for sample collections and assistance.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

**Geetha Annavi, PhD**
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

**Christina Yong Seok Yien, PhD**
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Member)

**Wan Norhamidah Binti Wan Ibrahim, PhD**
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Member)

**Simon Yung Wa Sin, PhD**
Assistant Professor
School of Biological Sciences
University of Hong Kong
(Member)

**ZALILAH MOHD SHARIFF, PhD**
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 10 September 2020

**Declaration by graduate student**

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____ Date: _____

Name and Matric No.: Nurul Adilah Binti Ismail, GS44635

**Declaration by Members of Supervisory Committee**

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: _____
Name of
Chairman of
Supervisory Committee: _____

Signature: _____
Name of
Member of
Supervisory Committee: _____

Signature: _____
Name of
Member of
Supervisory Committee: _____

Signature: _____
Name of
Member of
Supervisory Committee: _____

## TABLE OF CONTENTS


| | | Page |
|---|---|---|
| **ABSTRACT** | | i |
| ***ABSTRAK*** | | ii |
| **ACKNOWLEDGEMENTS** | | iv |
| **APPROVAL** | | v |
| **DECLARATION** | | vii |
| **LIST OF TABLES** | | xi |
| **LIST OF FIGURES** | | xiv |
| **LIST OF ABBREVIATIONS** | | xxiii |
| **CHAPTER** | | |
| **1** | **INTRODUCTION** | 1 |
| | 1.1 Objectives | 3 |
| **2** | **LITERATURE REVIEW** | 4 |
| | 2.1 Malayan Tapir | 4 |
| | 2.1.1 General characteristics | 4 |
| | 2.1.2 Evolution | 6 |
| | 2.1.3 Threats | 7 |
| | 2.2 Immune System in Vertebrates | 9 |
| | 2.3 Major Histocompatibility Complex (MHC) gene | 10 |
| | 2.4 MHC Gene Polymorphism | 14 |
| | 2.4.1 Heterozygote advantage | 16 |
| | 2.4.2 Pathogen-mediated selection | 16 |
| | 2.4.3 Sexual selection | 16 |
| | 2.5 MHC in conservation | 18 |
| | 2.6 Nucleotide Evolution | 19 |
| | 2.6.1 Substitutions in protein coding gene | 21 |
| | 2.6.2 Model selection for phylogenetic reconstruction | 23 |
| **3** | **METHODOLOGY** | 26 |
| | 3.1 Sample collection and DNA isolation | 26 |
| | 3.2 Primers Design | 27 |
| | 3.3 PCR Amplification, cloning and sequencing | 27 |
| | 3.4 Data Analaysis | 29 |
| | 3.4.1 Selection analysis | 29 |
| | 3.4.2 Phylogenetic analysis | 30 |
| | 3.4.3 Summary of software used | 30 |
| **4** | **RESULTS AND DISCUSSION** | 33 |
| | 4.1 gDNA and RNA isolation | 33 |

4.2 Primer design 36
4.3 PCR amplification, cloning and sequencing 50
4.4 Orthologous sequence comparison 66
4.4.1 MHC Class I 66
4.4.2 MHC Class II 69
4.5 Selection analysis: rate of synonymous and non-synonymous substitution 74
4.5.1 MHC Class I 74
4.5.2 MHC Class II 78
4.6 Phylogenetic analysis 83
4.6.1 MHC Class I 83
4.6.2 MHC Class II 84

**5 CONCLUSION AND RECOMMENDATION** 93
5.1 General Conclusion 93
5.2 Recommendation 94

**REFERENCES** 96
**APPENDICES** 109
**BIODATA OF STUDENT** 146
**PUBLICATION** 147

## LIST OF TABLES


| Table | | Page |
|---|---|---|
| 2.1 | Rates of synonymous and non-synonymous nucleotide substitution in different genes. Table adapted from Graur and Li (2000). Data from Li (1993). | 14 |
| 2.2 | Rate of synonymous and non-synonymous nucleotide substitution in class I and II gene in different organisms. | 15 |
| 2.3 | MHC-mate choice related studies in several wild population across different taxa. Table adapted from Huchard, Knapp, Wang, Raymond, and Cowlishaw (2010). M=Male F=Female | 17 |
| 2.4 | Pathogen effect on population with reduced genome-wide genetic variation. Table adapted from Radwan et al. (2010). | 19 |
| 2.5 | Several evolutionary models with their assumptions on base frequencies and substitution. The models are listed from least to more complex. | 25 |
| 3.1 | Summary of software used and their functions in this study. Software is listed in order of analysis conducted. | 32 |
| 4.1 | gDNA and RNA concentration value for isolation from seven Malayan tapirs individual. Tapir 1 sample was not used for RNA isolation. | 33 |
| 4.2 | Primer checking result for GC content, hairpin and temperature value | 46 |
| 4.3 | All working primers used for MHC Class I and II amplification. F = forward, R = reverse, Ta = annealing temperature, bp = base pair, td = touchdown. | 48 |
| 4.4 | MHC sequences identified from seven individuals of Malayan tapir. MHC class I sequences are divided into exon 2 and exon 3 region which code for alpha 1 and alpha 2 domains of the MHC class I molecule. MHC class II sequences consist of exon 2 region from two alpha 1 gene (DRA and DQA) and two beta 1 gene (DRB and DQB). Exon 2 of class II genes encode for alpha 1 and beta 1 domains of the molecules. | 55 |
| 4.5 | Presence of MHC class I alleles in the seven Malayan tapirs in this study. The individuals are denoted as T1-T7 (Tapir 1 -Tapir 7). | 62 |
| 4.6 | Presence of MHC class II alleles in the seven Malayan tapirs in this study. The individuals are denoted as T1-T7 (Tapir 1 – Tapir 7). | 62 |
| 4.7 | Sequence polymorphism of MHC Class I and II genes. | 63 |
| 4.8 | Percentage of similarities among class I exon 2 Malayan tapir sequences. | 64 |
| 4.9 | Percentage of similarities among class I exon 3 Malayan tapir sequences. | 65 |

| | | |
|---|---|---|
| 4.10 | Percentage of similarities of class II DRB and DQB sequences in Malayan tapir. | 66 |
| 4.11 | Rates of non-synonymous (dN) and synonymous (ds) substitutions for antigen-binding site (ABS) and non-ABS, and combined (ABS and non-ABS) at the Malayan tapir Tapirus indicus MHC class I loci as determined in PAML. ω indicate ratio of non-synonymous to synonymous nucleotide substitution. | 75 |
| 4.12 | Null and alternative models comparison determined in PAML for class I gene in Malayan tapir. | 76 |
| 4.13 | Rates of non-synonymous (dN) and synonymous (ds) substitutions for antigen-binding site (ABS) and non-ABS, and combined (ABS and non-ABS) at the Malayan tapir Tapirus indicus MHC class II loci determined in PAML. ω indicate ratio of non-synonymous to synonymous nucleotide substitution. | 79 |
| 4.14 | Null and alternative models comparison determined in PAML for class II DRB gene in Malayan tapir. | 81 |
| 4.15 | Positive selected sites were identified in models M2a and M8 by Bayes Empirical Bayes (BEB) Posterior Probabilities. | 82 |
| A1 | Pairwise distance exon 2 class I of Malayan tapir homologous gene to other species. Accession numbers for other species are listed in the result. The values is listed in matrix form. The lower the value shows shorter distance between two sequences and therefore the higher similarity of the seqeuences. | 129 |
| A2 | Pairwise distance exon3 class I of Malayan tapir homologous gene to other species. Accession numbers for other species are listed in the result. The values is listed in matrix form. The lower the value shows shorter distance between two sequences and therefore the higher similarity of the seqeuences. | 131 |
| A3 | Pairwise distance DQA class II of Malayan tapir homologous gene to other species. Accession numbers for other species are listed in the result. The values is listed in matrix form. The lower the value shows shorter distance between two sequences and therefore the higher similarity of the seqeuences. | 133 |
| A4 | Pairwise distance DQB class II of Malayan tapir homologous gene to other species. Accession numbers for other species are listed in the result. The values is listed in matrix form. The lower the value shows shorter distance between two sequences and therefore the higher similarity of the seqeuences. | 135 |
| A5 | Pairwise distance DRA class II of Malayan tapir homologous gene to other species. Accession numbers for other species are listed in the result. The values is listed in matrix form. The lower the value shows shorter distance between two sequences and therefore the higher similarity of the seqeuences. | 136 |

| | | |
|---|---|---|
| A6 | Pairwise distance DRB class II of Malayan tapir homologous gene to other species. Accession numbers for other species are listed in the result. The values is listed in matrix form. The lower the value shows shorter distance between two sequences and therefore the higher similarity of the seqeuences. | 138 |
| A7 | AIC value for Malayan tapir MHC class I exon 2 alignments with other species for phylogenetic analysis model selection. Model with lowest AIC value was selected for phylogenetic construction. | 140 |
| A8 | AIC value for Malayan tapir MHC class I exon 3 alignments with other species for phylogenetic analysis model selection. Model with lowest AIC value was selected for phylogenetic construction. | 140 |
| A9 | AIC value for Malayan tapir MHC class II DQA alignments with other species for phylogenetic analysis model selection. Model with lowest AIC value was selected for phylogenetic construction. | 141 |
| A10 | AIC value for Malayan tapir MHC class II DQB alignments with other species for phylogenetic analysis model selection. Model with lowest AIC value was selected for phylogenetic construction | 141 |
| A11 | AIC value for Malayan tapir MHC class II DRA alignments with other species for phylogenetic analysis model selection. Model with lowest AIC value was selected for phylogenetic construction. | 142 |
| A12 | AIC value for Malayan tapir MHC class II DRB alignments with other species for phylogenetic analysis model selection. Model with lowest AIC value was selected for phylogenetic construction | 142 |
| A13 | AIC value for class MHC class II alpha domain (DRA and DQA) alignments with other species for phylogenetic analysis model selection. Model with lowest AIC value was selected for phylogenetic construction. | 143 |
| A14 | AIC value for Malayan tapir MHC class II beta domain (DRB and DQB) alignments with other species for phylogenetic analysis model selection. Model with lowest AIC value was selected for phylogenetic construction. | 143 |

## LIST OF FIGURES


| Figure | | Page |
|---|---|---|
| 2.1 | Adult malayan tapir (*Tapirus indicus*). Malayan tapir exhibits white and black appearance. Source: IUCN database (https://www.iucnredlist.org/species/21472/45173636) | 4 |
| 2.2 | Malayan tapir distribution in Southeast Asia countries. The Malayan Tapir Range from Myanmar, Thailand, Peninsular Malaysia and Sumatera, Indonesia. The population distribution is highlighted in yellow shade on the made. Source: IUCN database (https://www.iucnredlist.org/species/21472/45173636) | 5 |
| 2.3 | Phylogenetic tree represent divergence of tapir species with equids and rhinoceros in the family perissodactyla. Figure adapted from Houck et al. (2000) and Steiner and Ryder (2011). | 7 |
| 2.4 | Genetic position of MHC locus in human and mouse. Human has eight class I locus (e.g. A, B and C) and six class II loci (e.g. DP, DQ, and DR). Mice has three class I loci (K, D and L) and 2 class II loci (A and E) (Penn & Potts, 1999). | 12 |
| 2.5 | Adaptive immune response assisted by MHC. (a) MHC class I receptors present endogenously derived antigens to cytotoxic T-cells. (b and c) MHC class II receptors are expressed on antigen presenting cells and present exogenously derived antigens to helper T-cells (Penn & Potts, 1999). | 13 |
| 2.6 | Jukes and Cantor one-parameter model. The rate of substitution in all direction is equal and denoted by α (Graur & Li, 2000). | 20 |
| 2.7 | Kimura two-parameter model of nucleotide substitutions. The rate of transition (α) may not be equal to transversion (β) (Graur & Li, 2000). | 21 |
| 4.1 | Agorose gel (2%) visualaisation of gDNA extraction from seven Malayan tapirs following manufacturer's protocol. The first lane on each gel indicate the ladder used and numbered lanes 1 -7 were loaded with gDNA extraction samples Tapir 1 – Tapir 7. | 34 |
| 4.2 | Agarose gel (2%) visualisation of RNA extraction from seven individuals using manufacturer's protocol. The first lane was loaded with ladder and numbered lanes 2 -6 were loaded with RNA samples extracted from the six Malayan tapirs (Tapir 2- Tapir 7). Lane 7 was empty. | 35 |
| 4.3 | Figure shows the alignment of the exon 2 MHC class I region from human and closely related species to tapirs (accession numbers are detailed in methodology). The suitable forward and reverse primers were designed within | 38 |

the highlighted region of the alignment from the 5' and 3' regions of the aligned sequences respectively.
4.4 Figure shows the alignment of the exon 3 MHC class I region from human and closely related species to tapirs (accession numbers are detailed in methodology). The suitable forward and reverse primers were designed within the highlighted region of the alignment from the 5' and 3' regions of the aligned sequences respectively. 40
4.5 Figure shows the alignment of the exon 2 MHC class II DRA region from closely related species to tapirs (accession numbers are detailed in methodology). The suitable forward and reverse primers were designed within the highlighted region of the alignment from the 5' and 3' regions of the aligned sequences respectively. 42
4.6 Figure shows the alignment of the exon 2 MHC class II DQB region from closely related species to tapirs (accession numbers are detailed in methodology). The suitable forward and reverse primers were designed within the highlighted region of the alignment from the 5' and 3' regions of the aligned sequences respectively. 44
4.7 Agarose gel (2%) visualisation of PCR amplifications. Figure a) shows PCR amplification of MHC class I region using designed primers using Malayan tapir samples. Figure b) shows the colony PCR done from the colonies formed by cloning the successful amplified a) samples. The first lanes on each were loaded with ladder while the subsequent lanes were loaded with PCR amplified samples. 50
4.8 Agarose gel (2%) visualisation of PCR amplifications. Figure a) shows PCR amplification of MHC class II DRA region using designed primers using Malayan tapir samples. Figure b) shows the colony PCR done from the colonies formed by cloning the successful amplified a) samples. The first lanes on each were loaded with ladder while the subsequent lanes were loaded with PCR amplified samples. 51
4.9 Agarose gel (2%) visualisation of PCR amplifications. Figure a) shows PCR amplification of MHC class II DRB region using published primers using Malayan tapir samples. Figure b) shows the colony PCR done from the colonies formed by cloning the successful amplified a) samples. The first lanes on each were loaded with ladder while the subsequent lanes were loaded with PCR amplified samples. 52
4.10 Agarose gel (2%) visualisation of PCR amplifications. Figure a) shows PCR amplification of MHC class II DQA region using published primers using Malayan tapir samples. Figure b) shows the colony PCR done from the colonies formed by cloning the successful amplified a) samples. The first lanes on each were loaded with ladder 53

while the subsequent lanes were loaded with PCR amplified samples.

4.11 Agarose gel (2%) visualisation of PCR amplifications. Figure a) shows PCR amplification of MHC class II DQB region using designed primers using Malayan tapir samples. Figure b) shows the colony PCR done from the colonies formed by cloning the successful amplified a) samples. The first lanes on each were loaded with ladder while the subsequent lanes were loaded with PCR amplified samples. 54

4.12 Amino acid sequence identity for the Malayan tapir *Tapirus indicus* class I exon 2 clones, rhinoceros (*Diceros bicornis*), equids (*Equus caballus*), human (*Homo sapiens*) and cattle (*Bos taurus*). The GenBank accession numbers for α1 (exon2) sequences from other mammals are AF055346 (Diceros_bicornis_DibiUA01), DQ083407 (Equus_caballus_Eqca100101) and DQ145597 (Equus_caballus_Eqca100201), GU812295 (Homo sapiens_HLA_A01010101) and L02834 (Bos_taurus_classIBoLA). *Numbers above the sequence* indicate the codon position in the α1 domain. *Single letters* and *dots* represent amino acids that are distinct from or identical to Tapirus indicus sequence for this domain respectively. *Dashes* indicate missing sequences. Putative ABSs were defined according to Reche and Reinherz (2003) and are marked with asterisk mark above the sequence. 67

4.13 Amino acid sequence identity for the Malayan tapir *Tapirus indicus* class I exon 3 clones, rhinoceros (*Rhinoceros unicornis Diceros bicornis,* and *Ceratotherum simum*), equids *(Equus caballus)*, and cattle (*Bos taurus*). The GenBank accession numbers for α2 (exon3) sequences from other mammals are AJ133670 (R_unicornis_classI), AJ055348 (D_bicornis_DibiUB02), XM_014795072 (Ceratotherium_simum_classi), DQ083407 (E_caballus_Eqca100101) and DQ083408 (E_caballus_Eqca200101), and L02834 (Bos_taurus_classI_BoLA). *Numbers above the sequence* indicate the codon position in the α2 domain. *Single letters* and *dots* represent amino acids that are distinct from or identical to *Tapirus indicus* sequence for this domain respectively. *Dashes* indicate missing sequences. Putative ABSs were defined according to Reche and Reinherz (2003) and are marked with asterisk mark above the sequence. 68

4.14 Amino acid sequence identity for the Malayan tapir *Tapirus indicus* class II exon 2 DRA clones, available Malayan tapir GenBank's sequence (*Tapirus indicus*), Baird's tapir (*Tapirus bairdii*), rhinoceros (*Diceros bicornis, Rhinoceros unicornis* and *Ceratotherum simum*), and equids (*Equus caballus*). The GenBank accession numbers for DRA 70

(exon2) sequences from other mammals are KM347953 (T_indicus_Tain-DRA-0104) and KM347956 (T_indicus-Tain-DRA-0106), AF113547 (T_bairdii_Taba-DRA-0101), AF113549 (D_bicornis_DRA-0101), AF113554 (R_unicornis_DRA-0501), AF113553 (Cera_simum_DRA-0401), and JQ254081 (E_caballus_Eqca-DRA-00102). *Numbers above the sequence* indicate the codon position in the DRα domain. *Single letters* and *dots* represent amino acids that are distinct from or identical to *Tapirus indicus* sequence for this domain respectively. *Dashes* indicate missing sequences. *Putative ABSs* were defined according to Reche and Reinherz (2003) and are marked with asterisk mark above the sequence.

4.15 Amino acid sequence identity for the Malayan tapir *Tapirus indicus* class II exon 2 DQA clones, human (*Homo sapiens*), equids (*Equus caballus*), bovine (*Bubalus bubalis*), boar (*Sus scrofa*), and coyote (*Canis latrans*). The GenBank accession numbers for DQA (exon2) sequences from other mammals are L3402 (HLA_DQA101011), JQ254060 (E_caballus_EqcaDQA100101), JQ254067 (E_caballus_EqcaDQA200202), KT428703 (B_bubalis_BubuDQA2103), AY285931 (Sus_scrofa_DQA1y) and AY126647 (Canis_latrans_DQA01701). *Numbers above the sequence* indicate the codon position in the DQα domain. *Single letters* and *dots* represent amino acids that are distinct from or identical to *Tapirus indicus* sequence for this domain respectively. *Dashes* indicate missing sequences. Putative ABSs were defined according to Reche and Reinherz (2003) and are marked with asterisk mark above the sequence. 71

4.16 Amino acid sequence identity for the Malayan tapir *Tapirus indicus* class II exon 2 DRB clones, equids (*Equus caballus*) and human (*Homo sapiens*). The GenBank accession numbers for DRB (exon2) sequences from other mammals are JQ254085 (E_caballus_DRB100101), JQ254084 (E_caballus_DRB100201), JQ254086 (E_caballus_DRB100301), and AF029288 (Homo_sapiens_DRB10101). *Numbers above the sequence* indicate the codon position in the DRβ domain. *Single letters* and *dots* represent amino acids that are distinct from or identical to *Tapirus indicus* sequence for this domain respectively. *Dashes* indicate missing sequences. Putative ABSs were defined according to Reche and Reinherz (2003) and are marked with asterisk mark above the sequence. 72

4.17 Amino acid sequence identity for the Malayan tapir *Tapirus indicus* class II exon 2 DQB clones, human (*Homo* sapiens), equids (*Equus caballus*), bovine (*Bos taurus*), boar (*Sus scrofa*) and dog (*Canis familiaris*). The GenBank 73

accession numbers for DQB (exon2) sequences from other mammals are L34101 (HLA_DQB050101), JQ254070 (E_caballus_EqcaDQB100101), JQ254069 (E_caballus_EqcaDQB100201), DQ093609 (Bos_taurus_BoLa_DQB), AY459300 (Sus_Scrofa_SLADQB1ax), and AF016905 (Canis_familiaris_DQB10010). *Numbers above the sequence* indicate the codon position in the DQβ domain. *Single letters* and *dots* represent amino acids that are distinct from or identical to *Tapirus indicus* sequence for this domain respectively. *Dashes* indicate missing sequences. Putative ABSs were defined according to Reche and Reinherz (2003) and are marked with asterisk mark above the sequence.

4.18 Phylogenetic tree of MHC class I α1 domain (exon2) sequences from Malayan tapir *Tapirus indicus*, equids, rhinoceros and other mammals including human and bovine. Bayesian posterior probabilities above 50% are shown above the branches. *Tapirus indicus* sequences in this tree are T_indicus_exon2_SMC154, T_indicus_exon2-THC1515, T_indicus_MLC158, T_indicus_BYC157 and T_indicus_MLC1522. The outgroup for this tree is M_rufogriseus_MaruUA01. 84

4.19 Phylogenetic tree of MHC class I α2 domain (exon3) sequences from Malayan tapir *Tapirus indicus*, equids, rhinoceros and other mammals including human and bovine. Bayesian posterior probabilities above 50% are shown above the branches. *Tapirus indicus* sequences in this tree are T_indicus_exon3_SMC154, T_indicus_exon3-BDC153, T_indicus_exon3, and T_indicus_exon3_MLC156. The outgroup for this tree is M_rufogriseus_MaruUA01. 85

4.20 Phylogenetic tree of MHC class II DRα domain (exon2) sequences from Malayan tapir *Tapirus indicus*, equids, rhinoceros and other mammals including human and fish (outgroup). Bayesian posterior probabilities above 50% are shown above the branches. *Tapirus indicus* sequences from this study in this tree are T.indicus_LR and T.indicus_PR. The outgroup in this tree is A_platyrthynchos_DRA. 87

4.21 Phylogenetic tree of MHC class II DQα domain (exon 2) sequences from Malayan tapir *Tapirus indicus*, equids, rhinoceros and other mammals including human bovine and canine. Bayesian posterior probabilities above 50% are shown above the branches. *Tapirus indicus* sequences in this tree are T.indicusDQA1 and T.indicusDQA2. The outgroup in this tree is canis_latrans_DQA01701. 88

4.22 Phylogenetic tree of MHC class II DRβ domain (exon 2) sequences from Malayan tapir *Tapirus indicus*, equids and other mammals including human and bovine. Bayesian 90

posterior probabilities above 50% are shown above the branches. *Tapirus indicus* sequences in this tree are T_indicus_BDDRbB32, T_indicus_LRDRB33, and T_indicus_THDRB1. The outgroup in this tree is F_passerinus_DRB.

4.23 Phylogenetic tree of MHC class II DQβ domain (exon 2) sequences from Malayan tapir *Tapirus indicus*, equids, rhinoceros and other mammals including human and canine. Bayesian posterior probabilities above 50% are shown above the branches. *Tapirus indicus* sequences in this tree are T_indicus_LRdqb3, T_indicus_bydqb9, and T_indicus_MLdqb5. The outgroup for this tree is canis_familiaris_DQb100101. 91

A1 Figure above shows letter of approved permit for blood collection from six Malayan tapir with restricted conditions. The permit was approved by Ministry of Natural Resources and Environment (NRE) through Department of Wildlife and National Park (DWNP) in 2016. 109

A2 Sequence confirmation for Malayan tapir MHC class I exon 2 isolated sequence SMC154 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 110

A3 Sequence confirmation for Malayan tapir MHC class I exon 2 isolated sequence THC1515 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 111

A4 Sequence confirmation for Malayan tapir MHC class I exon 2 isolated sequence MLC158 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 112

A5 Sequence confirmation for Malayan tapir MHC class I exon 2 isolated sequence BYC157 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and 113

accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity.

A6 Sequence confirmation for Malayan tapir MHC class I exon 2 isolated sequence MLC1522 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 114

A7 Sequence confirmation for Malayan tapir MHC class I exon 3 isolated sequence BDC153 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 115

A8 Sequence confirmation for Malayan tapir MHC class I exon 3 isolated sequence MLC158 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 116

A9 Sequence confirmation for Malayan tapir MHC class I exon 3 isolated sequence MLC156 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 117

A10 Sequence confirmation for Malayan tapir MHC class I exon 3 isolated sequence SMC154 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 118

A11 Sequence confirmation for Malayan tapir MHC class II DQA exon 2 isolated sequence BYDQA using BLAST program. The figure shows list of sequences most similar 119

to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity.

A12 Sequence confirmation for Malayan tapir MHC class II DQA exon 2 isolated sequence LRDQA using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 120

A13 Sequence confirmation for Malayan tapir MHC class II DQB exon 2 isolated sequence MLdqb5 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 121

A14 Sequence confirmation for Malayan tapir MHC class II DQB exon 2 isolated sequence LRdqb3 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 122

A15 Sequence confirmation for Malayan tapir MHC class II DQB exon 2 isolated sequence bydqb9 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 123

A16 Sequence confirmation for Malayan tapir MHC class II DRA exon 2 isolated sequence LR9 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. 124

| | | |
|---|---|---|
| A17 | Sequence confirmation for Malayan tapir MHC class II DRA exon 2 isolated sequence PR6 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. | 125 |
| A18 | Sequence confirmation for Malayan tapir MHC class II DRB exon 2 isolated sequence TuahDRB7 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. | 126 |
| A19 | Sequence confirmation for Malayan tapir MHC class II DRB exon 2 isolated sequence BDE32 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. | 127 |
| A20 | Sequence confirmation for Malayan tapir MHC class II DRB exon 2 isolated sequence LRE33 using BLAST program. The figure shows list of sequences most similar to the Malayan tapir sequence with the highest similarity at the top of the list. The figure also max score, total score, query cover, calculated e-value, percentage of identification and accession numbers of the sequence list. Generally, e-value lower than 1 is considered strong association of similarity. | 128 |
| A21 | Phylogenetic tree of MHC class II α domain (exon 2 of DRA and DQA) sequences from Malayan tapir *Tapirus indicus*, equids, rhinoceros and other mammals including human and bovine. Bayesian posterior probabilities above 50% are shown above the branches. | 144 |
| A22 | Phylogenetic tree of MHC class II β domain (exon 2 of DRB and DQB) sequences from Malayan tapir *Tapirus indicus*, equids, rhinoceros and other mammals including human and bovine. Bayesian posterior probabilities above 50% are shown above the branches. | 145 |

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ABS | Antigen Binding Site |
| AIC | Akaike information criterion |
| BEB | Bayes Empirical Bayes |
| BIC | Bayesian information criterion |
| bp | Base pairs |
| $d_N$ | Rate of non-synonymous substitution |
| $d$s | Rate of synonymous substitution |
| F81 | Felsenstein model |
| GTR | General time reversible model |
| HKY | Hasegawa-Kishino-Yano model |
| HLA | Human Leukocyte Antigen |
| IUCN | International Union for Conservation of Nature |
| JC69 | Jukes and Cantor Model |
| K80 | Kimura two-parameter model |
| LRT | Likelihood ratio test |
| MCMC | Markov Chain Monte Carlo |
| MHC | Major Histocompatibility Complex |
| ML | Maximum Likelihood |
| NRE | Ministry of Natural Resource and Environment |
| PAML | Phylogenetic Analysis by Maximum Likelihood |
| PAUP | Phylogenetic Analysis Using Parsimony |
| PBR | Peptide Binding Region |
| PERHILITAN | Jabatan Perlindungan Hidupan Liar dan Taman Negara |
| PFR | Permanent Forest Reserve |
| PP | Posterior Probability |
| PSS | Positive selection site |
| T1-T7 | Tapir 1 -Tapir 7 |
| $T_a$ | Annealing temperature |
| td | touchdown |
| Δ | Delta |

# CHAPTER 1

## INTRODUCTION

The Major Histocompatibility Complex (MHC) gene is a multigene family in vertebrates that plays important roles in adaptive immune system (Kimura, 1980; Klein, 1986; Janeway, Travers, Walport, & Shlomchik, 2001).This gene plays essential role in encoding cell surface glycoprotein known as MHC molecules. The MHC molecules are encoded by two major classes of the MHC gene which are class I and II genes. For example, in human class I genes are HL-A, HLA-B, and HLA-C while the class II genes are DR, DP, DQ ( Penn, 2002; Blum, Wearsch, & Cresswell, 2013).

When a cell is invaded with foreign pathogens such as viruses and bacteria, both class I and II MHC molecules will bind and present fragments of the foreign pathogen peptides on the cell's surface to T-cells and B-cells (Alberts et al., 2013). The T cells and B cells, once activated will initiate immediate immune response such as lysis of the infected cells. Perhaps due to its important function to recognise wide range of pathogens, the MHC gene is the most polymorphic genes in vertebrates (Janeway, Travers, Walport, & Shlomchik, 2001). The high diversity in the gene, particularly at the MHC gene that encode the peptide binding region (PBR) of the MHC molecules, is attributed to balancing selection, a type of selection that maintains the high allelic frequency and nucleotide diversity in the population (Hughes & Hughes, 1995; Llaurens, Whibley, & Joron, 2017; Koenig et al., 2019).

In most protein coding gene where selection is neutral, the rate of synonymous nucleotide substitutions (substitution that does not result in change in amino acid) is greater than the non-synonymous nucleotide substitutions (substitutions than results in changes in amino acid). This is because, non-synonymous nucleotide substitutions tend to change the amino acid and therefore are likely to be deleterious (Graur & Li, 2000). However, MHC gene that encode for the PBR of MHC molecule, has been observed to display higher rate of non-synonymous nucleotide substitutions than synonymous nucleotide substitutions. The higher rates in this gene may signal that the allele carrying the MHC gene is under selection could be advantageous in the population ( Li, 1993; Brandt, César, Goudet, & Meyer, 2018). This scenario is common in the MHC gene that encode for different species across multiple vertebrates taxa such as European badger *M. meles* (Sin, Dugdale, Newman, Macdonald, & Burke, 2012b, 2012a), Koala *P. cinereus* (Cheng et al., 2018) and Spotted pardalote *P. punctatus* (Balasubramaniam, Mulder, Sunnucks, Pavlova, & Melville, 2017). Furthermore, individuals that are heterozygotes at the MHC gene are deemed to be advantageous to survival as their MHC molecules can recognise wide range of pathogens (Nelson et al., 2004; Osborne et al., 2015; Phillips et al., 2018). In

contrast, individuals with reduced MHC diversity would have limited ability to recognise range of pathogens.

Malayan tapir *(Tapirus indicus*) is one of the five tapir species that belongs to the family Tapiridae and Order Perissodactyla. Currently, this species is listed as an endangered species in the IUCN Red List due to multiple factors predominated by habitat loss and human disturbance. Present Malayan tapir worldwide population is estimated to be only around 2000-2500 individuals, with 1000-1700 residing in Malaysia forest, thus calls for more efforts to conserve this mammal (Traeholt et al., 2016).

Among existing conservation efforts for this species include captive breeding and wild population monitoring. However, one of the most challenging factor that make it harder to recover from the low population number is that Malayan tapir has slow reproduction rate whereby they can generally produce one calf every two years after a long gestation period (390-395 days) (Barongi, 1993). The small number of population further increases the risk of inbreeding, which could lead to inbreeding depression in the population (Sommer, 2005; Benton et al., 2018).

Inbreeding which characterised by the loss of genetic variability including at the MHC loci will therefore decrease their ability to fight pathogens and increase their susceptibility towards diseases (Hedrick & Miller, 1994; Keller & Waller, 2002; Spielman, Brook, & Frankham, 2004) .The loss of genetic variability could happen when there is lack of allelic exchange between isolated populations. As their ability to fight pathogens decreases, the risk of extinction increases as a result of low adaptive flexibility and increased vulnerability to disease (Parmar et al., 2017).

One of the unfortunate situations can be observed in the case of Tasmanian devil (Ujvari & Belov, 2011). The reduced MHC diversity in this species due to inbreeding causes the MHC class I gene to lose ability to recognise infectious tumour invading this species as foreign and therefore, causes the tumour to spread even more (Morris, Wright, Grueber, Hogg, & Belov, 2015). O'Brien (1985) also reported high mortality in inbred cheetah due to coronavirus-associated feline infectious peritonitis. However, it is not clear in this study if this is the effect of genome wide inbreeding or caused by observed reduction in MHC variation in this species.

Infectious disease is dangerous particularly in endangered species such as Malayan tapir because it can wipe out the entire species population (Lafferty & Gerber, 2002; Parmar et al., 2017). Hence, adaptive genetic variation is an essential part of evolution and of high importance for long-term of a species (Smith, Acevedo-Whitehouse, & Pedersen, 2009; Hedrick, 2012). While studies on neutral genetic variation such as microsatellite markers, provides information on demographic history of natural population, adaptive variation provides

evidence of selective processes that act on the gene as a result of Malayan tapir individuals' interaction with the environment and adaptive changes such as infectious disease (Sommer, 2005). As mentioned above, because MHC molecule interact with antigen to trigger immune response, it is important for the host to possess MHC genetic variations to confer disease resistance. Balancing selection that accounts for MHC gene diversity, also results in perseverance of individual alleles and strongly differentiated allelic lineages in mammals including the Malayan tapir (Sommer, 2005; Parmar et al., 2017). Therefore, MHC gene is essential marker to assess genetic variation related to disease resistance in host natural population. While extensive studies have been done on MHC gene in human and several vertebrates, there is no study that has focused on MHC gene in the endangered Malayan tapir. Therefore, characterisation of the gene would provide basic understanding on the MHC gene in Malayan tapir.

## 1.1 Objectives

This study aims to:

1) Characterise the MHC class I and II genes that encode for PBR in Malayan tapir
2) Investigate whether Malayan tapir MHC genes belongs to monophyletic group
3) Investigate if there is occurrence of trans-species polymorphism.

Characterisation of this gene in Malayan tapir will facilitate our understanding on the evolutionary process that influences the MHC gene diversity in this species. Essentially, this study will serve as basis for further studies on MHC variability, mate choice and pathogen resistance in Malayan tapir.

## REFERENCES


Abd-Elsalam, K. A. (2003). Bioinformatic tools and guideline for PCR primer design. *African Journal of Biotechnology*, *2*(5), 91–95.

Akaike, H. (1974). A new look at the statistical model identification. In *Selected Papers of Hirotugu Akaike* (pp. 215-222): Springer.

Alberts, B., Bray, D., Hopkin, K., Johnson, A. D., Lewis, J., Raff, M., … Walter, P. (2013). *Essential cell biology*. Garland Science.

Albright-Fraser, D. G., Reid, R., Gerber, V., & Bailey, E. (1996). Polymorphism of DRA among equids. *Immunogenetics, 43*(5), 315-317.

Alrowaithi, M. A. (2013). *The use of genetic materials, RNA, to determine the age of saliva stains*.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol, 215*(3), 403-410. doi:10.1016/S0022-2836(05)80360-2

Annavi, G., Dawson, D. A., Horsburgh, G. J., Greig, C., Dugdale, H. L., Newman, C., . . . Burke, T. (2011). Characterisation of twenty-one European badger (Meles meles) microsatellite loci facilitates the discrimination of second-order relatives. *Conservation genetics resources, 3*(3), 515-518.

Aranda IV, R., Dineen, S. M., Craig, R. L., Guerrieri, R. A., & Robertson, J. M. (2009). Comparison and evaluation of RNA quantification methods using viral, prokaryotic, and eukaryotic RNA over a 104 concentration range. *Analytical Biochemistry*, *387*(1), 122–127.

Arumugam, K. A., & Annavi, G. (2018). Captive Breeding of Threatened Mammals Native to Southeast Asia–A Review on their Ex-situ Management, Implication and Reintroduction Guidelines. *Annual Research and Review in Biology*, 1–16.

Arumugam, K. A., Luan, L. Q., Ibrahim, W. N. B. W., Toh, M. B. M., Tah, M., Buesching, C. D., & Annavi, G. (2018). Influence of Enclosure Conditions and Visitors on the Behavior of Captive Malayan Tapir (Tapirus indicus): Implications for Ex-Situ Management and Conservation. *International Journal of Scientific and Research Publications*, *8*(7). https://doi.org/10.29322/IJSRP.8.7.2018.p7906

Balasubramaniam, S., Mulder, R. A., Sunnucks, P., Pavlova, A., & Melville, J. (2017). MHC class II β exon 2 variation in pardalotes (Pardalotidae) is shaped by selection, recombination and gene conversion. *Immunogenetics*, *69*(2), 101–111.

Barongi, R. (1993). Husbandry and conservation of tapirs Tapirus spp. *International Zoo Yearbook*, *32*(1), 7–15.

Becker, L., Nieberg, C., Jahreis, K., & Peters, E. (2009). MHC class II variation in the endangered European mink Mustela lutreola (L. 1761)—consequences for species conservation. *Immunogenetics, 61*(4), 281-288.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2008). GenBank. *Nucleic acids research, 37*(suppl_1), D26-D31.

Benton, C. H., Delahay, R. J., Smith, F. A., Robertson, A., McDonald, R. A., Young, A. J., … Hodgson, D. (2018). Inbreeding intensifies sex-and age-dependent disease in a wild mammal. *Journal of Animal Ecology*, *87*(6), 1500–1511.

Bernáldez, V., Rodríguez, A., Rodríguez, M., Sánchez-Montero, L., & Córdoba, J. J. (2017). Evaluation of different RNA extraction methods of filamentous fungi in various food matrices. *LWT-Food Science and Technology*, *78*, 47–53.

Bernatchez, L., & Landry, C. (2003). MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of evolutionary biology, 16*(3), 363-377.

Blum, J. S., Wearsch, P. A., & Cresswell, P. (2013). Pathways of antigen processing. *Annu Rev Immunol*, *31*, 443–473. https://doi.org/10.1146/annurev-immunol-032712-095910

Bodmer, W. (1972). Evolutionary significance of the HL-A system. *Nature*, *237*(5351), 139.

Bonneaud, C., Chastel, O., Federici, P., Westerdahl, H., & Sorci, G. (2006). Complex Mhc-based mate choice in a wild passerine. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1590), 1111–1116.

Bonner, K., Siemieniuk, R. A., Boozary, A., Roberts, T., Fajardo, E., & Cohn, J. (2014). Expanding access to HIV viral load testing: A systematic review of RNA stability in EDTA tubes and PPT beyond current time and temperature thresholds. *PloS One*, *9*(12), e113813.

Borghans, J. A., Beltman, J. B., & De Boer, R. J. (2004). MHC polymorphism under host-pathogen coevolution. *Immunogenetics*, *55*(11), 732–739.

Bos, D. H., & Posada, D. (2005). Using models of nucleotide evolution to build phylogenetic trees. *Developmental and Comparative Immunology*, *29*(3), 211–227.

Brandt, D. Y., César, J., Goudet, J., & Meyer, D. (2018). The effect of balancing selection on population differentiation: A study with HLA genes. *G3: Genes, Genomes, Genetics*, *8*(8), 2805–2815.

Browne, L., & Karubian, J. (2018). Habitat loss and fragmentation reduce effective gene flow by disrupting seed dispersal in a neotropical palm. *Molecular Ecology*, *27*(15), 3055–3069.

Carbajal-Borges, J. P., Godínez-Gómez, O., & Mendoza, E. (2014). Density, abundance and activity patterns of the endangered Tapirus bairdii in one of its last strongholds in southern Mexico. *Tropical Conservation Science*, *7*(1), 100–114.

Cheng, Y., Polkinghorne, A., Gillett, A., Jones, E. A., O'Meally, D., Timms, P., & Belov, K. (2018). Characterisation of MHC class I genes in the koala. *Immunogenetics*, *70*(2), 125–133.

Dasgupta, D. (1993). An overview of artificial immune systems and their applications. In *Artificial immune systems and their applications* (pp. 3–21). Springer.

de Oliveira Andrade, L. J., de Melo, P. R. S., França, L. S., França, L. S., Malta, A. M. A., & Paraná, R. (2011). Human major histocompatibility complex and thyroid dysfunction in hepatitis C carries. *Revista de Ciências Médicas e Biológicas*, *10*(2), 179–186.

Desjardins, P., & Conklin, D. (2010). NanoDrop microvolume quantitation of nucleic acids. *JoVE (Journal of Visualized Experiments)*, (45), e2565.

Dieffenbach, C., Lowe, T., & Dveksler, G. (1993). General concepts for PCR primer design. *PCR Methods Appl*, *3*(3), S30–S37.

Doherty, P. C., & Zinkernagel, R. M. (1975). Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, *256*(5512), 50.

Duarte, J. M. B., & González, S. (2010). *Neotropical cervidology: Biology and medicine of Latin American deer*. Funep Jaboticabal.

Duncan, M. (2018). Perissodactyls. In K. A. Terio, D. McAloose, & J. S. Leger (Eds.), *Pathology of Wildlife and Zoo Animals* (pp. 433-454): Academic Press.

Ejsmond, M. J., Phillips, K. P., Babik, W., & Radwan, J. (2018). The role of MHC supertypes in promoting trans-species polymorphism remains an open question. *Nature Communications*, *9*(1), 4362.

Ellis, S., Martin, A., Holmes, E., & Morrison, W. (1995). At least four MHC class I genes are transcribed in the horse: phylogenetic analysis suggests an unusual evolutionary history for the MHC in this species. *International Journal of Immunogenetics, 22*(3), 249-260.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evoutionl, 17*(6), 368-376.

Fletcher, W., & Yang, Z. (2010). The effect of insertions, deletions, and alignment errors on the branch-site test of positive selection. *Molecular Biology and Evolution*, *27*(10), 2257–2267.

Forsgren, E., Locke, B., Semberg, E., Laugen, A. T., & de Miranda, J. R. (2017). Sample preservation, transport and processing strategies for honeybee RNA extraction: Influence on RNA yield, quality, target quantification and data normalization. *Journal of Virological Methods*, *246*, 81–89.

Fraser, D. G., & Bailey, E. (1998). Polymorphism and multiple loci for the horse DQA gene. *Immunogenetics*, *47*(6), 487–490.

Gensterblum-Miller, E., Wu, W., & Sawalha, A. H. (2018). Novel transcriptional activity and extensive allelic imbalance in the human MHC region. *The Journal of Immunology*, *200*(4), 1496–1503.

Giese, A. R., & Hedrick, P. W. (2003). *Genetic variation and resistance to a bacterial infection in the endangered Gila topminnow*. *6*, 369–377. Cambridge University Press.

Gillingham, M. A., Courtiol, A., Teixeira, M., Galan, M., Bechet, A., & Cezilly, F. (2016). Evidence of gene orthology and trans-species polymorphism, but not of parallel evolution, despite high levels of concerted evolution in the major histocompatibility complex of flamingo species. *Journal of Evolutionary Biology*, *29*(2), 438–454.

Goldman, N., & Yang, Z. (1994). A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Molecular Biology and Evolution*, *11*(5), 725–736.

Graur, D., & Li, W. (2000). Fundamentals of molecular evolution. Sinauer Associations. *Inc., Sunderland, MA*.

Grimsley, C., Mather, K. A., & Ober, C. (1998). HLA-H: a pseudogene with increased variation due to balancing selection at neighboring loci. *Molecular Biology and Evolution*, *15*(12), 1581–1588.

Hasegawa, M., Kishino, H., & Yano, T. (1988). Phylogenetic inference from DNA sequence data. *Statistical theory and data analysis II (K. Mutusita, ed.). Elsevier, Amsterdam*, 1-13.

Hedrick, P., & Miller, P. (1994). Rare alleles, MHC and captive breeding. In *Conservation genetics* (pp. 187–204). Springer.

Hedrick, P. W. (2012). What is the evidence for heterozygote advantage selection? *Trends in Ecology & Evolution*, *27*(12), 698–704.

Hedrick, P. W., Parker, K. M., Miller, E. L., & Miller, P. S. (1999). Major histocompatibility complex variation in the endangered Przewalski's horse. *Genetics*, *152*(4), 1701–1710.

Hess, C. M., & Edwards, S. V. (2002). The Evolution of the Major Histocompatibility Complex in Birds: Scaling up and taking a genomic approach to the major histo compatibilty complex (mhc) of birds reveals surprising departures from generalities found in mammals in both large-scale structure and the mechanisms shaping the evolution of the mhc. *Bioscience, 52*(5), 423-431.

Hoffmann, J. A., Kafatos, F. C., Janeway, C. A., & Ezekowitz, R. (1999). Phylogenetic perspectives in innate immunity. *Science*, *284*(5418), 1313–1318.

Holmes, E., & Ellis, S. (1999). Evolutionary history of MHC class I genes in the mammalian order Perissodactyla. *Journal of Molecular Evolution*, *49*(3), 316–324.

Houck, M., Kingswood, S., & Kumamoto, A. (2000). Comparative cytogenetics of tapirs, genus Tapirus (Perissodactyla, Tapiridae). *Cytogenetic and Genome Research*, *89*(1–2), 110–115.

How, M. J., & Zanker, J. M. (2014). Motion camouflage induced by zebra stripes. *Zoology*, *117*(3), 163–170.

Huchard, E., Knapp, L. A., Wang, J., Raymond, M., & Cowlishaw, G. (2010). MHC, mate choice and heterozygote advantage in a wild social primate. *Molecular Ecology*, *19*(12), 2545–2561.

Hudson, R. R., Kreitman, M., & Aguadé, M. (1987). A test of neutral molecular evolution based on nucleotide data. *Genetics*, *116*(1), 153–159.

Hughes, A. L., & Hughes, M. K. (1995). Natural selection on the peptide-binding regions of major histocompatibility complex molecules. *Immunogenetics*, *42*(4), 233–243. (7672817).

Hughes, Austin L, & Yeager, M. (1998). Natural selection at major histocompatibility complex loci of vertebrates. *Annual Review of Genetics*, *32*(1), 415–435.

Janeway, C. A., Travers, P., Walport, M., & Shlomchik, M. (2004). *Immunobiology: The immune system in health and disease*. Garland science.

Janeway, C. A., Travers, P., Walport, M., & Shlomchik, M. J. (2001). The major histocompatibility complex and its functions. In *Immunobiology: The Immune System in Health and Disease. 5th edition*. Garland Science.

Janis, C. (1976). The evolutionary strategy of the Equidae and the origins of rumen and cecal digestion. *Evolution*, *30*(4), 757–774.

Janis, C. (2008). An evolutionary history of browsing and grazing ungulates. In *The ecology of browsing and grazing* (pp. 21–45). Springer.

Janova, E., Matiasovic, J., Vahala, J., Vodicka, R., Van Dyk, E., & Horin, P. (2009). Polymorphism and selection in the major histocompatibility complex DRA and DQA genes in the family Equidae. *Immunogenetics*, *61*(7), 513–527.

Joomun, S. C., Hooker, J. J., & Collinson, M. E. (2008). Dental wear variation and implications for diet: An example from Eocene perissodactyls (Mammalia). *Palaeogeography, Palaeoclimatology, Palaeoecology*, *263*(3–4), 92–106.

Jukes, T. H., & Cantor, C. R. (1969). Evolution of protein molecules. *Mammalian Protein Metabolism*, *3*(21), 132.

Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, *17*(5), 230–241.

Kiemnec-Tyburczy, K., Richmond, J., Savage, A., Lips, K., & Zamudio, K. (2012). Genetic diversity of MHC class I loci in six non-model frogs is shaped by positive selection and gene duplication. *Heredity*, *109*(3), 146.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, *16*(2), 111–120.

Kimura, M. (1981). Estimation of evolutionary distances between homologous nucleotide sequences. *Proceedings of the National Academy of Sciences, 78*(1), 454-458.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol, 33*(7), 1870-1874. doi:10.1093/molbev/msw054

Kurtz, B. M., Singletary, L. B., Kelly, S. D., & Frampton, A. R. (2010). Equus caballus major histocompatibility complex class I is an entry receptor for equine herpesvirus type 1. *Journal of virology, 84*(18), 9027-9034.

Klein, J. (1986). *Natural history of the major histocompatibility complex*. Wiley.

Koenig, D., Hagmann, J., Li, R., Bemm, F., Slotte, T., Neuffer, B., … Weigel, D. (2019). Long-term balancing selection drives evolution of immunity genes in Capsella. *Elife*, *8*, e43606.

Korbie, D. J., & Mattick, J. S. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols*, *3*(9), 1452.

Lafferty, K. D., & Gerber, L. R. (2002). Good medicine for conservation biology: The intersection of epidemiology and conservation theory. *Conservation Biology*, *16*(3), 593–604.

Lanave, C., Preparata, G., Saccone, C., & Serio, G. (1984). A new method for calculating evolutionary substitution rates. *J Mol Evol, 20*(1), 86-93.

Landry, C., Garant, D., Duchesne, P., & Bernatchez, L. (2001). 'Good genes as heterozygosity': The major histocompatibility complex and mate choice in Atlantic salmon (Salmo salar). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1473), 1279–1285.

Leinders-Zufall, T., Brennan, P., Widmayer, P., Maul-Pavicic, A., Jäger, M., Li, X.-H., … Boehm, T. (2004). MHC class I peptides as chemosensory signals in the vomeronasal organ. *Science*, *306*(5698), 1033–1037.

Lenz, T. L. (2018). Adaptive value of novel MHC immune gene variants. *Proceedings of the National Academy of Sciences*, *115*(7), 1414–1416.

Lenz, T. L., Eizaguirre, C., Kalbe, M., & Milinski, M. (2013). Evaluating patterns of convergent evolution and trans-species polymorphism at MHC immunogenes in two sympatric stickleback species. *Evolution*, *67*(8), 2400–2412.

Li, W. H. (1993). Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J Mol Evol, 36*(1), 96-99.

Li, W. (1997). *Molecular evolution.* Sinauer associates incorporated.

Lim, Q. L., Ismail, A., Arumugam, R., Ng, W. L., Yong, C. S. Y., Ismail, A., . . . Annavi, G. (2018). A revisit to a low-cost method for the isolation of microsatellite markers: the case of the endangered Malayan tapir (Tapirus indicus). *BioRxiv*, 384651.

Lindtke, D., Lucek, K., Soria-Carrasco, V., Villoutreix, R., Farkas, T. E., Riesch, R., … Nosil, P. (2017). Long-term balancing selection on chromosomal variants associated with crypsis in a stick insect. *Molecular Ecology*, *26*(22), 6189–6205.

Lively, C. M., & Dybdahl, M. F. (2000). Parasite adaptation to locally common host genotypes. *Nature*, *405*(6787), 679.

Llaurens, V., Whibley, A., & Joron, M. (2017). Genetic architecture and balancing selection: The life and death of differentiated variants. *Molecular Ecology*, *26*(9), 2430–2448.

Medici, E. P., Mangini, P. R., & Fernandes-Santos, R. C. (2014). Health assessment of wild lowland tapir (Tapirus Terrestris) populations in the Atlantic forest and Pantanal biomes, Brazil (1996–2012). *Journal of Wildlife Diseases*, *50*(4), 817–828.

Medzhitov, R., & Janeway, C. A. (2002). Decoding the patterns of self and nonself by the innate immune system. *Science*, *296*(5566), 298–300.

Medzhitov, R., & Janeway Jr, C. A. (1998). *Innate immune recognition and control of adaptive immune responses*. *10*, 351–353. Elsevier.

Miller, H. C., Bowker-Wright, G., Kharkrang, M., & Ramstad, K. (2011). Characterisation of class II B MHC genes from a ratite bird, the little spotted kiwi (Apteryx owenii). *Immunogenetics, 63*(4), 223-233. doi:10.1007/s00251-010-0503-7

Miller, H. C., & Lambert, D. M. (2004). Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). *Mol Ecol, 13*(12), 3709-3721. doi:10.1111/j.1365-294X.2004.02368.x

Miller, D., Tallmadge, R. L., Binns, M., Zhu, B., Mohamoud, Y. A., Ahmed, A., … Antczak, D. F. (2017). Polymorphism at expressed DQ and DR loci in five common equine MHC haplotypes. *Immunogenetics*, *69*(3), 145–156.

Miyata, T., & Yasunaga, T. (1980). Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. *Journal of Molecular Evolution*, *16*(1), 23–36.

Morris, K. M., Wright, B., Grueber, C. E., Hogg, C., & Belov, K. (2015). Lack of genetic diversity across diverse immune genes in an endangered mammal, the Tasmanian devil (S arcophilus harrisii). *Molecular Ecology*, *24*(15), 3860–3872.

Mosmann, T. R., & Coffman, R. L. (1989). Heterogeneity of cytokine secretion patterns and functions of helper T cells. In *Advances in immunology* (Vol. 46, pp. 111–147). Elsevier.

Nei, M., & Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, *3*(5), 418–426.

Nei, M., & Rooney, A. P. (2005). Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet, 39*, 121-152. doi:10.1146/annurev.genet.39.073003.112240

Nelson, G. W., Martin, M. P., Gladman, D., Wade, J., Trowsdale, J., & Carrington, M. (2004). Cutting edge: Heterozygote advantage in autoimmune disease: Hierarchy of protection/susceptibility conferred by HLA and killer Ig-like receptor combinations in psoriatic arthritis. *The Journal of Immunology*, *173*(7), 4273–4276.

O'Brien, S. J., & Evermann, J. F. (1988). Interactive influence of infectious disease and genetic diversity in natural populations. *Trends in Ecology & Evolution*, *3*(10), 254–259.

Ohashi, Y., Creek, K. E., Pirisi, L., Kalus, R., & Young, S. R. (2004). RNA degradation in human breast tissue after surgical removal: A time-course study. *Experimental and Molecular Pathology*, *77*(2), 98–103.

Olsen, S. L. (2003). *Horses through time*. Rowman & Littlefield.

Osborne, A. J., Pearson, J., Negro, S. S., Chilvers, B. L., Kennedy, M. A., & Gemmell, N. J. (2015). Heterozygote advantage at MHC DRB may influence response to infectious disease epizootics. *Molecular Ecology*, *24*(7), 1419–1432.

Pagán, H. J., Ferrer, T., & O'Corry-Crowe, G. (2018). Positive selection in coding regions and motif duplication in regulatory regions of bottlenose dolphin MHC class II genes. *PloS One*, *13*(9), e0203450.

Parmar, D. R., Mitra, S., Bhadouriya, S., Rao, T., Kunteepuram, V., & Gaur, A. (2017). Characterization of major histocompatibility complex class I, and class II DRB loci of captive and wild Indian leopards (Panthera pardus fusca). *Genetica*, *145*(6), 541–558.

Pearman, P. B., & Garner, T. W. (2005). Susceptibility of Italian agile frog populations to an emerging strain of Ranavirus parallels population genetic diversity. *Ecology Letters*, *8*(4), 401–408.

Pedersen, A. B., Jones, K. E., Nunn, C. L., & Altizer, S. (2007). Infectious diseases and extinction risk in wild mammals. *Conservation Biology*, *21*(5), 1269–1279.

Pei, B., Sisu, C., Frankish, A., Howald, C., Habegger, L., Mu, X. J., … Diekhans, M. (2012). The GENCODE pseudogene resource. *Genome Biology*, *13*(9), R51.

Penn, D. J. (2002). The scent of genetic compatibility: Sexual selection and the major histocompatibility complex. *Ethology*, *108*(1), 1–21.

Penn, Dustin J, & Potts, W. K. (1999). The evolution of mating preferences and major histocompatibility complex genes. *The American Naturalist*, *153*(2), 145–164.

Phillips, K. P., Cable, J., Mohammed, R. S., Herdegen-Radwan, M., Raubic, J., Przesmycka, K. J., … Radwan, J. (2018). Immunogenetic novelty confers a selective advantage in host–pathogen coevolution. *Proceedings of the National Academy of Sciences*, *115*(7), 1552–1557.

Piertney, S., & Oliver, M. (2006). The evolutionary ecology of the major histocompatibility complex. *Heredity*, *96*(1), 7.

Posada, D., & Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics, 14*(9), 817-818.

Radwan, J., Biedrzycka, A., & Babik, W. (2010). Does reduced MHC diversity decrease viability of vertebrate populations? *Biological Conservation*, *143*(3), 537–544.

Reche, P. A., & Reinherz, E. L. (2003). Sequence variability analysis of human class I and class II MHC molecules: Functional and structural correlates of amino acid polymorphisms. *Journal of Molecular Biology*, *331*(3), 623–641.

Robinson, J., Halliwell, J. A., Hayhurst, J. D., Flicek, P., Parham, P., & Marsh, S. G. (2014). The IPD and IMGT/HLA database: Allele variant databases. *Nucleic Acids Research*, *43*(D1), D423–D431.

Rodriguez, F., Oliver, J. L., Marin, A., & Medina, J. R. (1990). The general stochastic model of nucleotide substitution. *J Theor Biol, 142*(4), 485-501.

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics, 19*(12), 1572-1574.

Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X., & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics, 19*(18), 2496-2497.

Rychlik, W. (1993). Selection of primers for polymerase chain reaction. In *PCR Protocols* (pp. 31–40). Springer.

Schwensow, N., Castro-Prieto, A., Wachter, B., & Sommer, S. (2019). Immunological MHC supertypes and allelic expression: How low is the functional MHC diversity in free-ranging Namibian cheetahs? *Conservation Genetics*, *20*(1), 65–80.

Siddle, H. V., Kreiss, A., Eldridge, M. D., Noonan, E., Clarke, C. J., Pyecroft, S., … Belov, K. (2007). Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. *Proceedings of the National Academy of Sciences*, *104*(41), 16221–16226.

Sin, Y. W., Dugdale, H. L., Newman, C., Macdonald, D. W., & Burke, T. (2012a). Evolution of MHC class I genes in the European badger (Meles meles). *Ecology and Evolution*, *2*(7), 1644–1662.

Sin, Y. W., Dugdale, H. L., Newman, C., Macdonald, D. W., & Burke, T. (2012b). MHC class II genes in the European badger (Meles meles): Characterization, patterns of variation, and transcription analysis. *Immunogenetics*, *64*(4), 313–327.

Singh, V. K., Govindarajan, R., Naik, S., & Kumar, A. (2000). The effect of hairpin structure on PCR amplification efficiency. *Mol Biol Today*, *1*(3), 67–69.

Smith, K., Acevedo-Whitehouse, K., & Pedersen, A. (2009). The role of infectious diseases in biological conservation. *Animal Conservation*, *12*(1), 1–12.

Smith, K. F., Sax, D. F., & Lafferty, K. D. (2006). Evidence for the role of infectious disease in species extinction and endangerment. *Conservation Biology*, *20*(5), 1349–1357.

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology*, *2*(1), 16.

Spielman, D., Brook, B. W., & Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences*, *101*(42), 15261–15264.

Steiner, C. C., & Ryder, O. A. (2011). Molecular phylogeny and evolution of the Perissodactyla. *Zoological Journal of the Linnean Society*, *163*(4), 1289–1303.

Sun, J. C., Beilke, J. N., & Lanier, L. L. (2009). Adaptive immune features of natural killer cells. *Nature*, *457*(7229), 557.

Swofford, D. L., & Sullivan, J. (2003). Phylogeny inference based on parsimony and other methods using PAUP*. *The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny, cáp, 7*, 160-206.

Takuno, S., Nishio, T., Satta, Y., & Innan, H. (2008). Preservation of a pseudogene by gene conversion and diversifying selection. *Genetics*, *180*(1), 517–531.

Takahashi, K., Rooney, A. P., & Nei, M. (2000). Origins and divergence times of mammalian class II MHC gene clusters. *J Hered, 91*(3), 198-204.

Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol, 10*(3), 512-526. doi:10.1093/oxfordjournals.molbev.a040023

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution, 28*(10), 2731-2739.

Tao, N., Richardson, R., Bruno, W., & Kuiken, C. (2008). FindModel. *Avail-able from: http://www. hiv. lanl. gov/content/sequence/find-model/findmodel. html (date last accessed June 22, 2009)*.

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on mathematics in the life sciences, 17*(2), 57-86.

Tallmadge, R. L., Campbell, J. A., Miller, D. C., & Antczak, D. F. (2010). Analysis of MHC class I genes across horse MHC haplotypes. *Immunogenetics*, *62*(3), 159–172.

Tallmadge, R. L., Lear, T. L., & Antczak, D. F. (2005). Genomic characterization of MHC class I genes of the horse. *Immunogenetics*, *57*(10), 763–774.

Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2003). Multiple sequence alignment using ClustalW and ClustalX. *Current protocols in bioinformatics*(1), 2.3. 1-2.3. 22.

Thong, L. M., Choi, H., Kwon, O., Kim, J., Kim, Y., Oh, J., … Park, C. (2011). Systematic analysis of swine leukocyte antigen-DRB1 nucleotide polymorphisms using genomic DNA-based high-resolution genotyping and identification of new alleles. *Tissue Antigens*, *77*(6), 572–583.

Total Forested Areas in Malaysia (1990-2014) Retrieved from http://www.nre.gov.my/en-my/Forestry/Pages/Statistics-Forest.aspx

Traeholt, C., Novarino, W., bin Saaban, S., Shwe, N. M., Lynam, A. J., Zainuddin, Z. Z., & bin Mohd, S. (2016). tapirus indicus. The IUCN Red List of Threatened Species 2016: e.T21472A45173636. http://www.iucnredlist.org/details/21472/0

Trifonov, V. A., Stanyon, R., Nesterenko, A. I., Fu, B., Perelman, P. L., O'Brien, P. C., … Robinson, T. J. (2008). Multidirectional cross-species painting illuminates the history of karyotypic evolution in Perissodactyla. *Chromosome Research*, *16*(1), 89–107.

Ujvari, B., & Belov, K. (2011). Major histocompatibility complex (MHC) markers in conservation biology. *International Journal of Molecular Sciences*, *12*(8), 5168–5186.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., & Leunissen, J. A. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res, 35*(Web Server issue), W71-74. doi:10.1093/nar/gkm306

Vallone, P. M., & Butler, J. M. (2004). AutoDimer: A screening tool for primer-dimer and hairpin structures. *Biotechniques*, *37*(2), 226–231.

van de Bildt, M. W., Kuiken, T., Visee, A. M., Lema, S., Fitzjohn, T. R., & Osterhaus, A. D. (2002). Distemper outbreak and its effect on African wild dog conservation. *Emerging Infectious Diseases*, *8*(2), 212.

Vivier, E., Tomasello, E., Baratin, M., Walzer, T., & Ugolini, S. (2008). Functions of natural killer cells. *Nature Immunology*, *9*(5), 503.

von Helversen, B., Schooler, L. J., & Czienskowski, U. (2013). Are stripes beneficial? Dazzle camouflage influences perceived speed and hit rates. *PLoS One*, *8*(4), e61173.

Winternitz, J., Minchey, S., Garamszegi, L. Z., Huang, S., Stephens, P., & Altizer, S. (2013). Sexual selection explains more functional variation in the mammalian major histocompatibility complex than parasitism. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1769), 20131605.

Yamazaki, K., Beauchamp, G. K., Kupniewski, D., Bard, J., Thomas, L., & Boyse, E. (1988). Familial imprinting determines H-2 selective mating preferences. *Science*, *240*(4857), 1331–1332.

Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*(8), 1586–1591.

Yang, Z., & Nielsen, R. (2000). Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Molecular Biology and Evolution*, *17*(1), 32–43.

Yang, Z., Wong, W. S., & Nielsen, R. (2005). Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution*, *22*(4), 1107–1118.

Yeager, M., & Hughes, A. L. (1999). Evolution of the mammalian MHC: natural selection, recombination, and convergent evolution. *Immunol Rev, 167*(1), 45-58.

## PUBLICATION

Ismail, N.A., Sin Y.W., and Annavi,G. 2016. Characterisation of Major Histocompatibility Complex (MHC) gene in Malayan Tapir (*Tapirus indicus*). Proceedings for the International Conference on Conservation and Sustainable Use of the Tropical Rainforests, 378-380. ISNB: 978-967-12081-5-1

**UNIVERSITI PUTRA MALAYSIA**

**STATUS CONFIRMATION FOR THESIS / PROJECT REPORT AND COPYRIGHT**

**ACADEMIC SESSION :** Second Semester 2019/2020

**TITLE OF THESIS / PROJECT REPORT :**

ISOLATION AND CHARACTERISATION OF MAJOR HISTOCOMPATIBILITY COMPLEX GENE CLASS I AND II IN MALAYAN TAPIR (*Tapirus indicus* Desmarest)

**NAME OF STUDENT:** NURUL ADILAH BINTI ISMAIL

I acknowledge that the copyright and other intellectual property in the thesis/project report belonged to Universiti Putra Malaysia and I agree to allow this thesis/project report to be placed at the library under the following terms:

1. This thesis/project report is the property of Universiti Putra Malaysia.
2. The library of Universiti Putra Malaysia has the right to make copies for educational purposes only.
3. The library of Universiti Putra Malaysia is allowed to make copies of this thesis for academic exchange.

I declare that this thesis is classified as :

*Please tick (√ )

| | | |
|---|---|---|
| ☐ | **CONFIDENTIAL** | (Contain confidential information under Official Secret Act 1972). |
| ☐ | **RESTRICTED** | (Contains restricted information as specified by the organization/institution where research was done). |
| ☐ | **OPEN ACCESS** | I agree that my thesis/project report to be published as hard copy or online open access. |

This thesis is submitted for :

☐ **PATENT** Embargo from_____ until _____
(date) (date)

**Approved by:**

_____ _____

(Signature of Student) (Signature of Chairman of Supervisory Committee)
New IC No/ Passport No.: Name:

Date : Date :

**[Note : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization/institution with period and reasons for confidentially or restricted. ]**