Pertanika J. Trop. Agric. Sci. 40 (3): 329 - 342 (2017)

PERTANIKA JOURNALS

TROPICAL AGRICULTURAL SCIENCE

Journal homepage: http://www.pertanika.upm.edu.my/

Review Article

A Glimpse into the Genomic Outlook of the Long-Tailed Macaque (*Macaca fascicularis*)

Joey Ee Uli[1]*, Christina Yong Seok Yien[2] and Tan Soon Guan[1]

[1]Department of Cell and Molecular Biology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia
[2]Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ABSTRACT

The long-tailed macaque (*Macaca fascicularis*) is commonly used for biomedical researches. However, genetic variation within a population or among populations can significantly affect phenotypical outcomes of treatments tested on model organisms. As such, it is important for studies involving model organisms originating from different, or even the same geographical locations to have sufficient genomic and transcriptomic background of the model organisms that is used. This paper summarises the utilisation of next-generation sequencing (NGS) technology to sequence genomes and transcriptomes of long-tailed macaques from various geographical populations in general and the Malaysian long-tailed macaque in particular, and its importance in the context of population genetic studies.

Keywords: *Macaca fascicularis*, next-generation sequencing, genomics, genetic variation, transcriptomics, population genetics, biomedical science

INTRODUCTION

The long-tailed macaque *(Macaca fascicularis*) is currently regarded as the most heavily-utilised non-human primate (NHP) model in biomedical studies (Ogawa & Vallender, 2014). The macaque's close evolutionary relationship with humans enables the macaque to recapitulate human biology, physiology, behaviour, and

ARTICLE INFO

Article history:
Received: 27 August 2016
Accepted: 27 March 2017

E-mail addresses:
joeyeeuli@gmail.com (Joey Ee Uli),
chrisyong@upm.edu.my (Christina Yong Seok Yien),
sgtan_98@yahoo.com (Tan Soon Guan)
* Corresponding author

ISSN: 1511-3701

symptoms when presented with certain perturbations in the environment. Thus making it useful for researchers studying the various effects of drugs and treatments in a cellular environment. A wide array of studies have utilised the long-tailed macaque as NHP models (Nunamaker et al., 2013; Lee et al., 2014b; Silverstein et al., 2014; Berry et al., 2015; Gardner et al., 2015), and their efficacy in translational studies will continue to attract researchers in the foreseeable future. That being said, thought must be given to the genetic variation of long-tailed macaques originating from different geographical locations. Variations in the genome causes phenotypical changes that is observed through a wider spectrum of the model organism's reactions towards drug treatments. Hence, researchers need to take into consideration the functional variations contained within the genomes and transcriptomes of model NHPs before they embark on biomedical projects.

In recent years, genomics and transcriptomics play a pivotal role in biomedical studies involving NHPs. It is pertinent that genomic and transcriptomic information of model organisms is widely available to facilitate in experiment design. Next-generation sequencing (NGS) provides a platform that can expedite the process of obtaining the genomic and transcriptomic information of NHP organisms. In addition to its low cost and speed of generating sequences, NGS provides a huge volume of genomic and transcriptomic related useful for research. Various attempts have been made to sequence the genomes and transcriptomes of the long-tailed macaque from different geographical locations This paper will attempt to summarise the various genomic and transcriptomic sequencing endeavours carried out on the long-tailed macaques, as well as briefly outline phylogenetic and population genetic studies involving Malaysian long-tailed macaques. We will also discuss the need to sequence genomes and transcriptomes of the Malaysian long-tailed macaque in the context of population management.

GENOMES AND TRANSCRIPTOMES

Genomics and transcriptomics is the mainstay of biomedical studies which require the sequencing of whole genomes and/or transcriptomes, especially where non-model organisms are involved. Since the advent of NGS, whole genomes and transcriptomes can be easily sequenced.

To date, more than 3000 eukaryote genomes of various assembly levels have been deposited to NCBI's Genome database. Without having to rely on reference genomes, genomes can be assembled via de novo assembly to form contigs which subsequently are assembled into scaffolds. Newly assembled genomes are then annotated with gene information and their functions predicted by searching for orthologous regions in closely related species that are already annotated, in which case, established model organisms with their genomes completely sequenced and annotated. Insight into a genome is particularly useful for genomic comparative studies. The focus can either be on

evolutionary relationships between two or more species, or the genomic relatedness of several populations of organisms. Most noteworthy is how researchers no longer need to rely on inferring phylogeny or genetic diversity based merely on several genes, rather in the age of NGS, entire genomes can be used for higher resolution phylogeny inference.

A complementary approach to whole genome sequencing would be to sequence the transcriptome, as transcriptomes provide information pertaining to the quantification of gene expression, the discovery of novel transcripts and their isoforms, intragenic expression, as well as to understand modes of antisense regulation (Tarazona et al., 2011). RNA-seq is a remarkably cost-effective approach to producing massive amounts of data, capable of sequencing complete transcriptomes of almost any tissue, as well as measuring their levels of gene expression in a myriad of conditions (Ozsolak & Milos, 2011). The measurements of gene expression is based on the mapping and quantification of reads generated from the sequencing platform. In addition, RNA-seq provides a higher base pair resolution of gene expression measurement as opposed to conventional microarray techniques (Marioni et al., 2008), and also avoids the need for bacterial cloning of cDNA (Wang et al., 2009) – of which may cause sequences to be misrepresented or incomplete (Mortazavi et al., 2008). Unlike microarrays, which require prior knowledge of the genes of interest, one of RNA-seq's significant utility is the sequencing of novel transcriptomes. Hybridisations of novel transcripts to microarray probes designed for closely-related species may lead to unsuccessful hybridisations due to differences in the reference sequence and transcript in question. This makes it difficult to study non-model organisms which are more likely to not have their entire transcriptomes sequenced and catalogued. With RNA-seq, transcriptomes can be assembled without the need for a reference genome of the organism in question. By using paired-end sequencing reads, researchers can discover exon splice junctions and novel transcript isoforms within a transcriptome, which is not possible with the microarray platform.

WHOLE GENOME SEQUENCING OF THE LONG-TAILED MACAQUES

Following the completion of the human genome project in 2003, concerted efforts have been made to sequence model organisms seen beneficial to the biomedical field. Prior to the onset of next-generation sequencing, researchers have successfully sequenced the genome of an Indian rhesus macaque via whole-genome shotgun sequencing approach (Gibbs et al., 2007). Of major significance to the complete sequencing of the rhesus macaque's genome is its ability to provide insights into the biological pathways and the functionality of an adult rhesus macaque. Using the newly assembled draft rhesus macaque genome (rheMac2) Gibbs et al. (2007) and Osada et al. (2008) compared Philippine and Cambodian-Thailand hybrid *M. fascicularis* via sequencing of complementary DNA (cDNA) clones with

M. mulatta cDNA sequences predicted from the rheMac2 assembly. However, the team relied mostly on aligning the *M. mulatta* and *M. fascicularis* sequences with human genome sequences to identify orthologs, and thus estimate the genetic divergence between *M. fascicularis* and *M. mulatta*. Due to the rheMac2 genome assembly's lack of species-specific transcript annotation information, the presence of gaps, and the discovery of scaffold misassemblies (Norgren 2013), efforts were made to annotate the rheMac2 genome through RNA sequencing studies of the rhesus macaque (Pipes et al., 2013) as well as genome annotation endeavours (Zhang et al., 2013; Peng et al., 2014). A new rhesus macaque genome assembly was produced by Zimin et al. (2014) using NGS technology to correct the assembly errors in the rheMac2 assembly and the rhesus macaque genome assembly done by Yan et al. (2011). This assembly called MacaM, was built independently of the rheMac2 assembly and was also supplemented by RNA-seq data that was carried out simultaneously with the MacaM whole-genome sequencing. The drafting and annotation of the rhesus macaque genome opened doors for biomedical studies which rely y on the model organisms' gene expression patterns towards novel drug treatments.

One of the first whole-genome sequencing endeavour of the long-tailed macaque was accomplished by Ebeling et al. in 2011. They employed the NGS platform to sequence the genome of a Mauritian long-tailed macaque using the whole genome shotgun sequencing approach. The team assembled the long-tailed macaque genome by utilising the rheMac2 genome, as well as the human genome (hg18) based on homology. With the draft genome, the team carried out transcriptome predictions to develop *M. fascicularis*-specific gene expression microarrays. Their ultimate goal was to provide reliable genomic information of the long-tailed macaque for biomedical research and drug testing.

The genomes of a Vietnamese long-tailed macaque and a Chinese rhesus macaque were drafted by Yan et al. in 2011. They compared their genome sequences with that of the rheMac2 genome assembly and their findings suggest that the Vietnamese long-tailed macaque may have undergone introgression after hybridisation with the Chinese rhesus macaque. In addition, they also observed genetic differences in orthologues related to biomedical studies between the three macaque genomes they studied. The researchers postulated that macaque populations from distinct geographical locations are likely to have genetic variation.

In 2012, Higashino et al. endeavoured to construct the genome of the Malaysian long-tailed macaque by employing the resequencing method with reference to the rheMac2 genome assembly. Their rationale was that the Malaysian long-tailed macaque has shown to exhibit higher genetic diversity in contrast with their Vietnamese counterparts (Osada et al., 2010). The team's study supported Yan et al.'s (2011) postulation that the Vietnamese long-tailed

macaque had undergone introgression with the Chinese rhesus macaque.

These findings were further strengthened by Fan et al.'s (2014) assembly of the Tibetan macaque (*Macaca thibetana*) genome by conducting genetic diversity assessments among the genomes of the Tibetan macaque, the Chinese rhesus macaque, the Vietnamese long-tailed macaque, and the Malaysian long-tailed macaque. They found long-tailed macaque populations to display higher genetic diversity than populations of the rhesus macaques. Genome-wide comparisons also showed the Tibetan macaques were more closely related to the rhesus macaques compared to the long-tailed macaques. Subsequently, they also postulated that the Tibetan macaques underwent admixture with the Chinese rhesus macaque resulting from overlapping geographical distribution and mating habits.

The most recent whole-genome sequencing of the long-tailed macaque was undertaken by Osada et al. in 2015. Six Mauritian long-tailed macaque genomes were sequenced and it was determined that the overall nucleotide diversity of the Mauritian macaques was 23% smaller than the Malaysian long-tailed macaques. Further phylogenetic analysis between the newly sequenced Mauritian long-tailed macaque, Malaysian, Vietnamese, and Chinese long-tailed macaque suggested that the Mauritian macaques were genetically closer to the Malaysian macaques. Genome-wide comparisons between two or more closely related species are beneficial for genetic variation discovery studies, annotations of functional genes via orthologue comparisons, phylogenetic inference, and differential gene expression studies.

TRANSCRIPTOMIC SEQUENCING OF THE LONG-TAILED MACAQUE

Prior to NGS technology becoming ubiquitous, cloning and capillary sequencing were essential methodologies for sequencing the transcriptome of an organism. However, this approach involves the time consuming and costly effort of cloning an immense amount of cDNA sequences and sequencing all the cloned sequences. Despite the tedious work flow and cost, this was a method commonly used in sequencing expressed sequence tags (ESTs) to conduct transcript profiling. In 2005, Magness et al. sequenced the transcriptome of the rhesus macaque using this approach. They constructed the rhesus macaque transcriptome by mapping their reads to orthologs in the human messenger RNA (mRNA) RefSeq sequences as well as rhesus macaque mitochondrial sequences found in GenBank. Magness et al. also encountered rhesus macaque cDNA sequences that did not have putative orthologs in the human cDNA or genome, of which they were unable to determine their species specificity due to the then unavailability of a completed rhesus macaque genome. In addition, Magness et al. determined the percentage of sequence divergence between the rhesus macaque and humans in the coding and non-coding sequence levels, as well as in the amino acid level. Osada et al. (2008) also sequenced the transcripts of the long-

tailed and rhesus macaque using a similar approach of sequencing tens of thousands of cDNA clones.

Using RNA-seq technology, Huh et al. (2012) sequenced the transcriptome of long-tailed macaques and identified genes involved in various biological responses of the macaque. A total of 16 tissues originating from two macaque individuals obtained from Vietnam were used and 175 transcripts identified, 81 of which were experimentally validated. In an expansion of Huh et al.'s 2012 study, Park et al. (2013) used Huh et al.'s transcriptome sequencing data to select appropriate reference genes for real-time quantitative PCR (RT-qPCR) normalisation procedures involved in gene expression studies of the long-tailed macaque. This study provides a benchmark for research in gene expression studies on the long-tailed macaque. Subsequently, Lee et al. (2014b) and Park et al. (2015) used transcriptomic data from Huh et al. (2012) in two separate gene expression studies to investigate the quantitative gene expression of insulin/insulin-like growth factor, amyloid precursor protein, and tau-phosphorylation-related genes in Alzheimer's disease with the long-tailed macaque as a model organism.

A collaboration between researchers from many NPRCs, Pipes et al. (2013) set out to profile the transcriptomes of 13 NHP species of biomedical importance. Among the NHPs selected for the large-scale study include long-tailed macaques of Indochinese and Mauritian origin, the pig tailed macaque (*Macaca nemestrina*), rhesus macaques of Chinese and Indian origin, the Japanese macaque (*Macaca fuscata*), and other NHPs. The team employed multiple methods of library preparation to capture the deeper set of transcriptomic data, including coding transcripts, non-coding transcripts, and delineating information for strand-specific and anti-sense transcription detection. Parallel to the RNA-seq studies, the 'NHP Reference Transcriptome Resource' was established to deposit sequencing data as well as to serve as a public database and online community for researchers. Peng et al. (2014) further expanded the NHPRTR project by sequencing the same 15 tissues of 11 NHP species and subspecies – including the Indochinese and Mauritian long-tailed macaque. Their goal was to make tissue-specific RNA-seq data originating from the various NHPs sequenced in this endeavour. They also set out to improve the Indian rhesus macaque and Mauritian long-tailed macaque transcriptome annotations, filling up transcript sequence gaps in the genomes of the respective macaques, and providing novel isoforms for annotated genes and also unannotated intergenic transcripts enriched with noncoding RNA.

Lee et al. (2014a) sequenced the transcriptomes of the long-tailed macaque, African green monkey (*Chlorocebus aethiops*), and rhesus macaque using the Illumina GAIIx platform. They sequenced six tissue samples as well as blood samples harvested from the long-tailed macaque and proceeded with a reference-guided assembly of the long-tailed macaque by mapping the sequencing reads to the long-tailed macaque

genome. The team performed a comparative tissue-specific gene expression analysis of their long-tailed macaque and African green monkey transcriptomes with transcriptomes of other primates, including *Homo sapiens, Pan troglodytes, Gorilla gorilla*, and *Macaca mulatta*. Principal component analysis revealed the consistent gene expression profiles of orthologous protein-coding genes across the primates analysed. Furthermore, the team also identified and validated novel transcripts and splice isoforms obtained from their sequencing data. In addition to the transcriptome assembly, Lee et al. (2014a) also developed a Multi-Species Annotation (MSA) pipeline that enables systematic annotations of transcriptome assemblies by employing the use of BLAST and NCBI's nt database.

GENETIC VARIATIONS IN THE LONG-TAILED MACAQUES AND THE NEED FOR COMPREHENSIVE GENOMIC AND TRANSCRIPTOMIC INFORMATION

Genomic comparisons of the long-tailed macaques are mostly in concordance with genetic data. With genomic datasets, Higashino et al. (2012) and Fan et al. (2014) determined that the Malaysian long-tailed macaque has the highest genetic diversity among the populations of long-tailed macaques, a pattern that is similarly observed in Smith et al. (2007). Osada et al.'s (2015) findings using genomic data to determine the nucleotide diversity of Mauritian long-tailed macaques also mirror the findings of Smith et al. (2007). Genomic data (Yan et al., 2011; Higashino et al., 2012) also tally with findings of introgression between Vietnamese long-tailed macaques and Chinese rhesus macaques based on genetic data (Tosi et al., 2002; Bonhomme et al., 2009; Kanthaswamy et al., 2008). The parallel between genomic and genetic results shows the promise of genomics in population genetic studies. Genomics allows the expansion of the number of genes/ nucleotide characters utilised, providing for a larger sampling size for the construction and inference of phylogenetic trees, and single nucleotide polymorphism analysis. NGS vastly facilitates the acquisition of genomes and transcriptomes, which in turn provides the larger sample size for further downstream analyses at a fraction of the time, cost, and effort.

Due to inter- and intraspecific genetic variations in NHP models, a more robust and complete annotation of reference genomes from various locations and populations is needed in order to select populations with genetic backgrounds suitable for relevant experiments (Haus et al., 2014). Various studies employing macaques have observed different phenotypical reactions involving individuals from different geographical locations, or even among individuals from the same geographical origin. Trichel et al. (2002) observed differences in the progression of acquired immune deficiency syndrome (AIDS)-like viral infection in two different populations of rhesus macaques obtained from India and China, whereby the rhesus macaques from China lived longer than the Indian macaques. A study by Seekatz et al. (2013) into the

efficacy of Shigella vaccines administered to Mauritian and non-Mauritian long-tailed macaques showed that the two populations responded differently, arising from the presence of distinct gut microbiota in the Mauritian macaques. Genotyping efforts of TRIM5α by Zhang et al. (2016) revealed polymorphisms of the gene that suppresses the viral replication of HIV2-ROD in rhesus macaques, and suggests the need to genotype TRIM5α of rhesus macaques before commencing HIV studies that employ these model organisms. While NHP model organisms certainly assist in drug studies, it is essential that reference genomes of NHP models that originate from various locations are sequenced in order to design a solid research framework.

In a decade where Malaysia faces indiscriminate levelling of forests (Hamdan et al., 2016), the population of long-tailed macaques is becoming more fragmentised and threatened. Families of macaques are forced out of their habitats and into human dwellings where they risk of conflict with humans, capture and culling. Though the long-tailed macaques are not yet extinct their numbers are dwindling faster than initially expected (Eudey, 2008). Populations of macaques face the threat of inbreeding due to their fragmented habitats, which will inevitably reduce their genetic diversity overtime (Mona et al., 2014). Researchers can mitigate the possibility of population fragmentation through population management techniques. Following the introduction of high-throughput sequencing technology, whole genomes and transcriptomes can be more easily sequenced, assembled, and analysed. This approach is beneficial for population genetics researchers because large genomic and/or transcriptomic datasets can be generated more cheaply (Shendure & Ji, 2008; Metzker, 2010). From the whole genome or transcriptome dataset, and with appropriate pipelines – reviewed extensively in Ekblom & Galindo, (2011) and Davey et al., (2011) – researchers can easily extract thousands of sequences and design genetic markers for population genetics studies.

Presently, there is a need to sequence the genome and transcriptome of long-tailed macaques of Peninsular Malaysia and Borneo Malaysia. Out of the 50 *M. fascicularis* subspecies classified by Fooden (1995), *Macaca fascicularis* is the most widely distributed *M. fascicularis* subspecies in Peninsular and Borneo Malaysia (Groves 2001; Brandon-Jones et al., 2004). By utilising mitochondrial DNA (Tosi & Coke, 2007; Abdul-Latiff et al., 2014b), mitogenomic phylogeny (Liedigk et al., 2015), and examining Y-chromosome gene flow (Rovie-Ryan et al., 2013), researchers have observed phylogeography separations of Malaysian long-tailed macaque populations into mainland (Peninsular Malaysia) and insular (Borneo Malaysia) populations. Abdul-Latiff et al. (2014a) and Smith et al. (2007) have shown that the Malaysian population of long-tailed macaques are monophyletic with no shared haplotypes with other Southeast

Asian long-tailed macaque populations, with the Malaysian populations having higher levels of nucleotide diversity compared to the other long-tailed macaque populations of Southeast Asia. Higher genetic diversity results in a wider range of genotypic and phenotypic reactions to drug treatments (Osada et al., 2015), which could prove to be a valuable insight in biomedical research. In light of the long-tailed macaque's suitability as NHP models in drug-safety testing, and biomedical and disease-related studies, the importance of obtaining deep-sequencing genomic data of this species is crucial. A genomic and transcriptomic repository of not only long-tailed macaques, but other NHPs from various geographical locations is essential for sustainable and substantial biomedical, disease-related, and drug-safety testing research.

CONCLUSION

With the benefit of cutting costs, saving time, and the usage of relatively small amounts of tissue for the sequencing task, high throughput sequencing of the transcriptome can expedite the acquisition of genetic markers for the population genetic studies of the long-tailed macaques. A faster and more efficient management of the population of long-tailed macaques can hopefully be achieved, although this goes hand-in-hand with the many government and non-government agencies involved. Researchers have a duty to provide information needed for the respective parties to act.

REFERENCES

Abdul-Latiff, M. A. B., Ruslin, F., Faiq, H., Hairul, M. S., Rovie-Ryan, J. J., Abdul-Patah, P., … & Md-Zain, B. M. (2014a). Continental Monophyly and Molecular Divergence of Peninsular Malaysia's *Macaca fascicularis fascicularis*. *BioMed Research International, 2014*(2014), 1-18. doi: http://dx.doi.org/10.1155/2014/897682

Abdul-Latiff, M. A. B., Ruslin, F., Fui, V. V., Abu, M. H., Rovie-Ryan, J. J., Abdul-Patah, … & Md-Zain, B. M. (2014b). Phylogenetic Relationships of Malaysia's Long-Tailed Macaques, *Macaca fascicularis*, based on Cytochrome b Sequences. *ZooKeys*, (407), 121-139.

Berry, N., Ham, C., Alden, J., Clarke, S., Stebbings, R., Stott, J., … & Almond, N. (2015). Live Attenuated Simian Immunodeficiency Virus Vaccination Confers Superinfection Resistance against Macrophage-Tropic and Neurovirulent Wild-Type SIV Challenge. *Journal of General Virology, 96*(7), 1918-1929.

Bonhomme, M., Cuartero, S., Blancher, A., & Crouau-Roy, B. (2009). Assessing Natural Introgression in 2 Biomedical Model Species, the Rhesus Macaque (*Macaca mulatta*) and the Long-Tailed Macaque (*Macaca fascicularis*). *Journal of Heredity, 100*(2), 158-169.

Brandon-Jones, D., Eudey, A. A., Geissmann, T., Groves, C. P., Melnick, D. J., Morales, J. … & Stewart, C. B. (2004). Asian Primate Classification. *International Journal of Primatology, 25*(1), 97-164.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-Wide Genetic Marker Discovery and Genotyping using Next-Generation Sequencing. *Nature Reviews Genetics, 12*(7), 499-510.

Ebeling, M., Kung, E., See, A., Broger, C., Steiner, G., Berrera, M., … & Certa, U. (2011). Genome-Based Analysis of the Nonhuman Primate *Macaca fascicularis* as a Model for Drug Safety Assessment. *Genome Research, 21*(10), 1746-1756.

Ekblom, R., & Galindo, J. (2011). Applications of Next Generation Sequencing in Molecular Ecology of Non-Model Organisms. *Heredity, 107*(1), 1-15.

Eudey, A. A. (2008). The Crab-Eating Macaque (*Macaca fascicularis*): Widespread and Rapidly Declining. *Primate Conservation, 23*(1), 129-132.

Fan, Z., Zhao, G., Li, P., Osada, N., Xing, J., Yi, Y., … & Li, J. (2014). Whole-Genome Sequencing of Tibetan Macaque (*Macaca thibetana*) Provides New Insight into the Macaque Evolutionary History. *Molecular Biology and Evolution, 31*(6), 1475-1489. doi: 10.1093/molbev/msu104

Fooden, J. (1995). *Systematic Review of Southeast Asia Long-Tail Macaques, Macaca fascicularis (Raffles [1821])*. Chicago: Field Museum of Natural History.

Gardner, J., Rudd, P. A., Prow, N. A., Belarbi, E., Roques, P., Larcher, T., … & Suhrbier, A. (2015). Infectious Chikungunya Virus in the Saliva of Mice, Monkeys and Humans. *PLoS One, 10*(10), e0139481.

Gibbs, R. A., Rogers, J., Katze, M. G., Bumgar ner, R., Weinstock, G. M., Mardis, E. R., … & Zwieg, A. S. (2007). Evolutionary and Biomedical Insights from the Rhesus Macaque Genome. *Science, 316*(5822), 222-234.

Groves, C. P. (2001). *Primate taxonomy*. Washington: Smithsonian Institution Press.

Hamdan, O., Rahman, K. A., & Samsudin, M. (2016). Quantifying Rate of Deforestation and CO2 Emission in Peninsular Malaysia using Palsar Imageries. In, *8th IGRSM International Conference and Exhibition on Remote Sensing & GIS (IGRSM 2016)* (pp. 1-11). Bristol, United Kingdom: IOP Publishing.

Haus, T., Ferguson, B., Rogers, J., Doxiadis, G., Certa, U., Rose, N. J., … & Roos, C. (2014). Genome Typing of Nonhuman Primate Models: Implications for Biomedical Research. *Trends in Genetics, 30*(11), 482-487.

Higashino, A., Sakate, R., Kameoka, Y., Takahashi, I., Hirata, M., Tanuma, R., … & Osada, N. (2012). Whole-Genome Sequencing and Analysis of the Malaysian Cynomolgus Macaque (Macaca fascicularis) Genome. *Genome Biology, 13*(7), R58-R69. doi: 10.1186/gb-2012-13-7-r58

Huh, J. W., Kim, Y. H., Park, S. J., Kim, D. S., Lee, S. R., Kim, K. M., … & Chang, K. T. (2012). Large-Scale Transcriptome Sequencing and Gene Analyses in the Crab-Eating Macaque (*Macaca fascicularis*) for Biomedical Research. *BMC Genomics, 13*(1), 163-175. doi: 10.1186/1471-2164-13-163

Kanthaswamy, S., Satkoski, J., George, D., Kou, A., Erickson, B. J. A., & Smith, D. G. (2008). Hybridization and Stratification of Nuclear Genetic Variation in *Macaca mulatta* and *M. fascicularis*. *International Journal of Primatology, 29*(5), 1295-1311.

Lee, A., Khiabanian, H., Kugelman, J., Elliott, O., Nagle, E., Yu, G. Y., … & Rabadan, R. (2014a). Transcriptome Reconstruction and Annotation of Cynomolgus and African Green Monkey. *BMC Genomics, 15*(1), 846-857.

Lee, Y., Kim, Y. H., Park, S. J., Huh, J. W., Kim, S. H., Kim, S. U., … & Chang, K. T. (2014b). Insullin/ IGF Signaling-Related Gene Expression in the Brain of a Sporadic Alzheimer's Disease Monkey Model Induced by Intracerebroventricular Injection of Streptozotocin. *Journal of Alzheimer's Disease, 38*(2), 251-267.

Liedigk, R., Kolleck, J., Böker, K. O., Meijaard, E., Md-Zain, B. M., Abdul-Latiff, M. A., … & Roos, C. (2015). Mitogenomic Phylogeny of the Common Long-Tailed Macaque (*Macaca fascicularis fascicularis*). *BMC Genomics, 16*(1), 222-232. doi: 10.1186/s12864-015-1437-0

Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The Power and Promise of Population Genomics: from Genotyping to Genome Typing. *Nature Reviews Genetics, 4*(12), 981-994.

Magness, C. L., Fellin, P. C., Thomas, M. J., Korth, M. J., Agy, M. B., Proll, S. C., … & Iadonato, S. P. (2005). Analysis of the *Macaca mulatta* Transcriptome and the Sequence Divergence between Macaca and Human. *Genome Biology, 6*(7), R60-R75. doi: 10.1186/gb-2005-6-7-r60

Marioni, J., Mason, C., Mane, S., Stephens, M., & Gilad, Y. (2008). RNA-Seq: An Assessment of Technical Reproducibility and Comparison with Gene Expression Arrays. *Genome Research, 18*(9), 1509-1517.

Metzker, M. L. (2010). Sequencing Technologies - the Next Generation. *Nature Reviews Genetics, 11*(1), 31-46.

Mona, S., Ray, N., Arenas, M., & Excoffier, L. (2014). Genetic Consequences of Habitat Fragmentation during a Range Expansion. *Heredity, 112*(3), 291-299.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and Quantifying Mammalian Transcriptomes by RNA-Seq. *Nature Methods, 5*(7), 621-628.

Norgren Jr., R. B. (2013). Improving Genome Assemblies and Annotations for Nonhuman Primates. *ILAR Journal, 54*(2), 144-153.

Nunamaker, E. A., Halliday, L. C., Moody, D. E., Fang, W. B., Lindeblad, M., & Fortman, J. D. (2013). Pharmacokinetics of 2 Formulations of Buprenorphine in Macaques (*Macaca mulatta* and *Macaca fascicularis*). *Journal of the American Association for Laboratory Animal Science, 52*(1), 48-56.

Ogawa, L. M., & Vallender, E. J. (2014). Genetic Substructure in Cynomolgus Macaques (Macaca Fascicularis) on the Island of Mauritius. *BMC Genomics, 15*(1), 748-751. doi: 10.1186/1471-2164-15-748

Osada, N., Hashimoto, K., Kameoka, Y., Hirata, M., Tanuma, R., Uno, Y., … & Takahashi, I. (2008). Large-Scale Analysis of Macaca Fascicularis Transcripts and Inference of Genetic Divergence between *M. fascicularis* and *M. mulatta*. *BMC Genomics, 9*(1), 90-104. doi: 10.1186/1471-2164-9-90

Osada, N., Hettiarachchi, N., Babarinde, I. A., Saitou, N., & Blancher, A. (2015). Whole-Genome Sequencing of Six Mauritian Cynomolgus Macaques (*Macaca fascicularis*) reveals a Genome-Wide Pattern of Polymorphisms under Extreme Population Bottleneck. *Genome Biology and Evolution, 7*(3), 821-830.

Osada, N., Uno, Y., Mineta, K., Kameoka, Y., Takahashi, I., & Terao, K. (2010). Ancient Genome-Wide Admixture Extends beyond the Current Hybrid Zone between *Macaca fascicularis* and *M. mulatta*. *Molecular Ecology, 19*(14), 2884-2895.

Ozsolak, F., & Milos, P. M. (2011). RNA Sequencing: Advances, Challenges and Opportunities. *Nature Reviews Genetics, 12*(2), 87-98.

Park, S. J., Kim, Y. H., Huh, J. W., Lee, S. R., Kim, S. H., Kim, … & Chang, K. T. (2013). Selection of New Appropriate Reference Genes for RT-QPCR Analysis via Transcriptome Sequencing of Cynomolgus Monkeys (*Macaca fascicularis*). *PLoS One, 8*(4), e60758-e60765. doi: 10.1371/journal.pone.0060758

Park, S. J., Kim, Y. H., Nam, G. H., Choe, S. H., Lee, S. R., Kim, S. U., … & Chang, K. T. (2015). Quantitative Expression Analysis of APP Pathway and Tau Phosphorylation-Related Genes in the ICV STZ-Induced Non-Human Primate Model of Sporadic Alzheimer's Disease. *International Journal of Molecular Sciences, 16*(2), 2386-2402.

Peng, X., Thierry-Mieg, J., Thierry-Mieg, D., Nishida, A., Pipes, L., Bozinoski, M., … & Mason, C. E. (2014). Tissue-Specific Transcriptome Sequencing Analysis Expands the Non-Human Primate Reference Transcriptome Resource (NHPRTR). *Nucleic Acids Research, 43*(D1), D737-D742. doi: 10.1093/nar/gku1110

Pipes, L., Li, S., Bozinoski, M., Palermo, R., Peng, X., Blood, P., … & Katze, M. G. (2013). The Non-Human Primate Reference Transcriptome Resource (NHPRTR) for Comparative Functional Genomics. *Nucleic Acids Research, 41*(D1), D906-D914.

Rovie-Ryan, J. J., Abdullah, M. T., Sitam, F. T., Abidin, Z. Z., & Tan, S. G. (2013). Y-Chromosomal Gene Flow of *Macaca fascicularis* (Cercopithecidae) between the Insular and Mainland Peninsula of Penang State, Malaysia. *Journal of Science and Technology in the Tropics, 9*, 113-126.

Sandler, N. G., Bosinger, S. E., Estes, J. D., Zhu, R. T., Tharp, G. K., Boritz, E., … & Douek, D. C. (2014). Type I Interferon Responses in Rhesus Macaques Prevent SIV Infection and Slow Disease Progression. *Nature, 51*1(7511), 601-605.

Seekatz, A. M., Panda, A., Rasko, D. A., Toapanta, F. R., Eloe-Fadrosh, E. A., Khan, A. Q., … & Fraser, C. M. (2013). Differential Response of the Cynomolgus Macaque Gut Microbiota to Shigella Infection. *PLoS One, 8*(6), e64212-e64224. doi: 10.1371/journal.pone.0064212

Shendure, J., & Ji, H. (2008). Next-Generation DNA Sequencing. *Nature Biotechnology, 26*(10), 1135-1145.

Silverstein, M. G., Kaplan, J. R., Appt, S. E., Register, T. C., & Shively, C. A. (2014). Effect of Soy Isoflavones on Thyroid Hormones in Intact and Ovariectomized Cynomolgus Monkeys (*Macaca fascicularis*). *Menopause, 21*(10), 11361142.

Smith, D. G., Mcdonough, J. W., & George, D. A. (2007). Mitochondrial DNA Variation within and among Regional Populations of Longtail Macaques (*Macaca fascicularis*) in Relation to other Species of the Fascicularis Group of Macaques. *American Journal of Primatology, 69*(2), 182-198.

Tarazona, S., García-Alcalde, F., Dopazo, J., Ferrer, A., & Conesa, A. (2011). Differential Expression in RNA-Seq: A Matter of Depth. *Genome Research, 2*1(12), 2213-2223.

Tosi, A. J., & Coke, C. S. (2007). Comparative Phylogenetics Offer New Insights into the Biogeographic History of *Macaca fascicularis* and the Origin of the Mauritian Macaques. *Molecular Phylogenetics and Evolution, 42*(2), 498-504.

Tosi, A. J., Morales, J. C., & Melnick, D. J. (2002). Y-Chromosome and Mitochondrial Markers in *Macaca fascicularis* Indicate Introgression with Indochinese M. Mulatta and a Biogeographic Barrier in the Isthmus of Kra. *International Journal of Primatology, 23*(1), 161-178.

Trichel, A. M., Rajakumar, P. A., & Murphey-Corb, M. (2002). Species-Specific Variation in SIV Disease Progression between Chinese and Indian Subspecies of Rhesus Macaque. Journal of *Medical Primatology, 31*(4-5), 171-178.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: A Revolutionary Tool for Transcriptomics. *Nature Reviews Genetics, 10*(1), 16-19.

Yan, G., Zhang, G., Fang, X., Zhang, Y., Li, C., Ling, F., … & Wang, Z. (2011). Genome Sequencing and Comparison of Nonhuman Primate Animals Models, the Cynomolgus and Chinese Rhesus Macaques. *Nature Biotechnology, 29*(11), 1019-1023.

Zhang, H. L., Liu, F. L., Jin, Y. B., Deng, Q., Liu, B. L., Zhuo, M., … & Ling, F. (2016). The Effects of TRIM5α Polymorphism on HIV-2ROD and SIVmac239 Replication in PBMCs from Chinese Rhesus Macaques and Vietnamese-Origin Cynomolgus macaques. *Virology, 487*, 222-229.

Zhang, S. J., Liu, C. J., Shi, M., Kong, L., Chen, J. Y., Zhou, W. Z., … & Li, C. Y. (2013). Rhesusbase: A Knowledgebase for the Monkey Research Community. *Nucleic Acids Research, 41*(D1), D892-D905. doi: 10.1093/nar/gks835

Zimin, A. V., Cornish, A. S., Maudhoo, M. D., Gibbs, R. M., Zhang, X., Pandey, S., … & Norgren Jr, R. B. (2014). A New Rhesus Macaque Assembly and Annotation for Next-Generation Sequencing Analyses. *Biology Direct, 9*(1), 20-34. doi: 10.1186/1745-6150-9-20