



## Genetic Analysis of Saltwater Crocodile (*Crocodylus porosus*) from Sarawak River Basin, Sarawak using Cytochrome Oxidase I and Cytochrome b Gene Analysis

Reymathi a/p Nadarajan<sup>1</sup>, Ruhana Hassan<sup>1,2</sup>, Mohd Izwan Zulaini Abdul Gani<sup>3</sup>, Mohd Azuwan Hassan<sup>1</sup>, Rambli Ahmad<sup>3</sup>, Besar Ketol<sup>2</sup>

Received: 16-09-2022

Accepted: 16-11-2022

### Abstract

*Crocodylus porosus* or known as saltwater crocodile is the common crocodylian species in Sarawak. *C. porosus* is currently placed under Appendix II of Convention on International Trade in Endangered Species of Wild Flora and Fauna (CITES) in Sarawak, thus trade is allowed with Permits from the authority. Sarawak River Basin (SRB) supports a large *C. porosus* wild population and will be involved in regulated crocodile harvesting exercise. However, there is limited information on the genetics of crocodiles in the area. Thus, this study is designed to determine the genetic diversity and molecular phylogeny of *C. porosus* using combine gene of Cytochrome Oxidase I (COI) and Cytochrome b (Cyt b) mtDNA gene analysis. Scutes and tissues samples were collected from wild crocodiles in SRB, later subjected to standard molecular biology techniques and DNA sequencing. Approximately 499 bp of COI gene and 418 bp of Cyt b genes from 9 individuals had been successfully sequenced. Nucleotide composition analysis showed that Cytosine is the highest while the lowest was Guanine and higher proportion of A+T bases was observed compared to G+C bases for both genes. The average nucleotide frequencies combine genes are as follow: Adenine = 28.5%, Thymine = 26.8%, Cytosine = 28.7% and Guanine = 16.0%, respectively. In this study, intraspecific genetic divergence values of *C. porosus* ranged between 0 to 1.7% for both genes whereas interspecific genetic divergence values range between 8.5 to 10%. Based on both COI and Cyt b gene information, phylogenetics trees constructed using Neighbour-joining, Maximum Parsimony and Maximum Likelihood methods show that *C. porosus* is monophyletic, with *Tomistoma schlegelii* as the outgroup. This is an ongoing project as more samples are needed to produce a large database for future metagenomics study of Sarawak crocodiles.

**Keywords:** crocodile, genetics, Cytochrome oxidase I, Cytochrome b.

1-Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak.

2-Centre for Pre-University Studies, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak

3-Sarawak Forestry Corporation, Lot 218, KCLD, Jalan Tapang, Kota Sentosa, 93250 Kuching, Sarawak

\*Corresponding author email: 20020014@siswa.unimas.my

## Introduction

Crocodylia is an order of mostly large, predatory, semiaquatic reptiles, live in various aquatic habitats such as rivers, marshes, swamps, forest streams and elbow lakes. They can be found in over 90 of the world's countries and islands (Martin, 2008). Out of 24 species of crocodiles in the world, only two species can be found in Sarawak namely the saltwater crocodile, *Crocodylus porosus* and the Malayan false gharial, *Tomistoma schlegellii* (Hassan et al., 2016, 2018). Both species can be easily distinguished as both crocodiles have a distinctive snout feature. The false gharials have elongated and slender snouts compared to those of the saltwater crocodiles, which are shorter and blunt (Grigg and Gans, 1993). In Sarawak, both species occupy different habitats; *C. porosus* mainly in waterway areas near to the coast, meanwhile the *T. schlegellii* prefers inland freshwater rivers with peat swamps habitat and black water (Hassan et al., 2016). *C. porosus* could be found in almost all river basins in Sarawak and frequently involve in human-crocodile conflicts (Abdul-Gani, 2019; Tisen and Ahmad, 2010).

Majority of saltwater crocodiles in Sarawak exist in the wild environment, well- distributed in the twenty-one river basins concentrating in the mangrove areas along the coastal zone (Hassan et al. 2018; Abdul-Gani, 2019). Continuous conservation effort in Sarawak has been successful as reflected in terms of crocodile densities in the rivers during in crocodile surveys

over the years (Sarawak Forestry Corporation, SFC) thus in 2016, Convention on the International Trade of Endangered Species (CITES) had agreed to transfer crocodile in Sarawak from Appendix I to Appendix II, allowing harvest of the wild population for the socio-economic benefit of the people living along the rivers (Hassan et al, 2018).

Cytochrome oxidase subunit 1 (CO1) gene is a widely used genetic marker due to its conserved characteristics, making it suitable for DNA barcoding exercise in animal species. Besides, this gene is highly variable among species and it has higher evolutionary rate compared to nuclear genes, making it useful to distinguish closely related species (Roldan et al. 2009). Similar to COI gene, Cytochrome b (Cyt b) gene is a protein coding gene found in the mitochondria of eukaryotic cells. Cyt b is commonly used to determine phylogenetic relationships between organisms, within family and genera, due to its sequence variability.

Meganathan et al. (2010) had sequenced a complete mitochondrial genome of *C. porosus*. This information was later combined with other gene information and concluded that their phylogenetic study support the well-established intergeneric crocodylian phylogenetic relationships. Mc Aliley et al (2006) had analysed two nuclear genes and two mitochondrial genes and they managed to support the hypothesis that *C. cataphractus* is not a member of *Crocodylus* or *Osteolaemus*. Preliminary molecular work using Cyt b and 12S rRNA gene were unable to resolve relationships among crocodiles from different areas in Sarawak by Shoon (2009) probably due to small number of samples used. On the other hand, Abdullah (2010) had used combination of Cyt b and COI genes analysis

and managed to partially resolve phylogenetic relationships among samples obtained from different rivers in Sarawak. For *Tomistoma*, Cyt b - control region mtDNA gene analysis, they revealed that the genetic diversity among *T. schlegelii* ranging from 0.16% to 3.34%, with unique haplotypes in Sumatera as well as in Northern Sarawak; thus each *T. schlegelii* population deserved its own management strategies to ensure the conservation of the species (Md Adzhar & Hassan, 2017).

Although Sarawak supports a high population of crocodiles, limited studies on genetic of crocodiles had been carried out, thus understanding of the population at the genetic level is still at its infancy stage. This study aims to sequence and analyze COI gene and Cyt b gene from crocodiles in Sarawak River Basins. Findings from this study could be used in formulating sustainable utilization of crocodiles in the rivers, for the benefits of local people.

### Materials and Methods

Total of 9 samples of crocodiles (*C. porosus*) were collected in SRB (Table 1.0) with the help by a group of experienced personals for large adult crocodiles and for baby crocodiles were caught manually by using scoop net (Abdul-Gani, 2014). Each sample were given voucher number according to the place. Tissues or scutes of crocodiles were collected for DNA work were kept in preservation tube contain 70% ethanol solutions and all samples were

transferred in -20°C freezer until further DNA analysis in Aquatic Molecular Laboratory, Universiti Malaysia Sarawak.

**Table1.0: Voucher numbers for samples according to sampling area.**

Voucher number	Sampling Area	Number of samples (n)
BK	Bako	1
BS	Samarahan	2
HI	Sungai Mang Hilir	1
HU	Sungai Mang Hulu	1
SM	Sungai Maranek	2
SR	Saribas	1
SY	Sebuyao	1
<b>TOTAL</b>		<b>9</b>

Total genomic DNA was extracted from crocodile samples (tissues and scutes) using 2% modified cetyltrimethylammonium bromide (CTAB) protocol (Doyle and Doyle, 1987) and followed by 1% Agarose Gel Electrophoresis (AGE).

PCR was conducted using thermocycler (MyCycler™) to amplify COI gene using with forward primer COI<sub>f</sub> (5'- CCT GCA GGA GGA GGA GAY CC-3'), and reverse primer COI<sub>a</sub> (5'-AGT ATA AGC GTC TGG GTA GTC-3'). As for Cyt b gene, forward primer GLU (5'-TGA TAT GAA AAA CCA TCG TTG-3'), and reverse primer CB2 (5'-CCC TCA GAA TGA TAT TTG TCC TCA-3') were used (Palumbi et al., 2002). Standard PCR with total volume of 25µl was prepared for each batch together with a negative control. Thermal profile used are as follow: preliminary denaturation at 94°C for 15 seconds, annealing at 58°C (COI gene), 45°C (Cyt gene) for 30 seconds and extension at 72°C for 50 seconds and followed by a final 4 minutes extension step at 72°C. Once the amplification completed, all the PCR products were subjected to AGE with 100bp DNA ladder. All positive results were then sent to commercial company for PCR product

cleaning and single pass DNA sequencing.

The raw DNA sequences observe the quality of the chromatogram (CHROMAS PRO). Then, the updated DNA sequences were subjected to nucleotide Basic Local Alignment Search Tool (BLAST) to find the similarities and confirmed the species (Altschul et al., 1990). All sequences were uploaded in MEGA X for multiple alignment, followed by calculation of genetic distance using Tamura Nei-parameter model and construction of phylogenetic trees (Kumar et al., 2016). Three phylogenetic trees were constructed namely: Maximum Parsimony (MP), Neighbour-Joining (NJ) and Maximum likelihood (ML)

used Bootstrap Method (1000 value). Based on the best-fit substitution model in Mega X, Hasegawa-Kishino-Yahon model and gamma distributed with invariant sites (G + I) were used to construct ML while Tamura- Nei model and gamma distributed were used to construct NJ and MP.

## Results

Approximately 499 bp of COI gene and 418 bp of Cyt b genes from 9 individuals had been successfully sequenced. Table 2 shows the BLAST results, all COI and Cyt b gene sequences from this study matched with two *C. porosus* DNA sequence from NCBI Genbank database, Accession No.: DQ273698.1 and NC008143.1 with average of 96.94% to 100% similarity. The E-value for all samples matched Genbank data were zero.

**Table 2.0: Basic Local Alignment Search Tool (BLAST) result of *C. porosus* samples in Sarawak with *C. porosus* Genbank data.**

Species/ Accession no.	Genebank data ( <i>Crocodylus porosus</i> )			
	DQ273698.1		NC008143.1	
	E-value	Identities (%)	E-value	Identities (%)
BK001	0	100	0	99.69
BS001	0	98.29	0	97.98
BS002	0	99.84	0	99.53
HI001	0	97.11	0	96.94
HU001	0	99.06	0	98.75
SM001	0	98.73	0	98.57
SM002	0	99.61	0	99.41
SR001	0	99.06	0	98.75
SY001	0	99.61	0	99.42

The complete sequences alignment of nine *C. porosus* samples with additional sequences from the Genbank were aligned together using Mega X software. List of data obtained from gene bank is in Table 2. The length of final alignment was 917bp. From a total of 917 aligned bases, 706 sites (77%) were conserved while 211 sites (23%) were variable. From the variables sites,

75 sites (8.2%) were singleton sites and 136 (14.8%) were parsimonious informative sites.

The intraspecific genetic distance for *C. porosus* samples from this study and other countries were in the range between 0-1.7% (Table 4.0). The interspecific genetic distance for *C. porosus* with all other species are as follow: values of genetic distance of *C. porosus* and *C. intermedius* were 9.3-10.5%; *C. porosus* and *C. mindorensis* were 8.5-

10.5%, *C. porosus* and *C. palustris*, the values were 9.1-10.6%.

Table 3 shows the nucleotide base composition of nine *C. porosus* based on combination of COI and Cyt b gene. The average of the nucleotide base composition is as follow: thymine (T) = 26.8%, cytosine (C) = 28.7%, adenine

(A) = 28.5% and guanine (G) = 16%. Cytosine showed the highest values of nucleotide composition and guanine showed the lowest values which the range from 28.7% to 28.8% and 15.9% to 16.0%, respectively. Higher proportion of A+T bases were observed compared to G+C bases which the values are 55.3% and 44.7%, respectively.

**Table 3.0: The average of nucleotide base composition of *C. porosus* samples (Sarawak) with the combination gene of COI and cyt b**

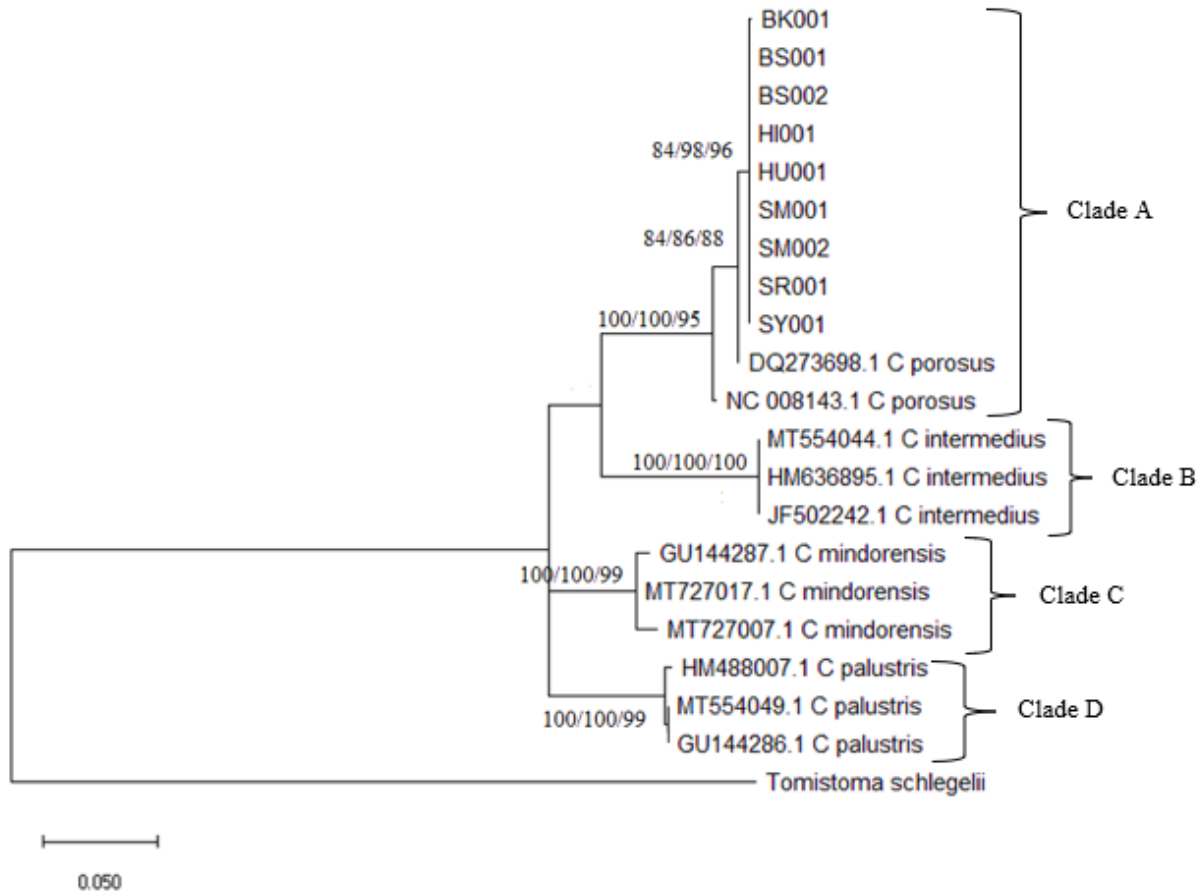
Sample	Nucleotide base			
	T(U)	C	A	G
BK001	26.9	28.8	28.4	15.9
BS001	26.8	28.7	28.5	16.0
BS002	26.8	28.7	28.5	16.0
HI001	26.8	28.7	28.5	16.0
HU001	26.8	28.7	28.5	16.0
SM001	26.8	28.7	28.5	16.0
SM002	26.8	28.7	28.5	16.0
SR001	26.8	28.7	28.5	16.0
SY001	26.8	28.7	28.5	16.0
Average	26.8	28.7	28.5	16.0

**Table 4.0: Pairwise distances analysis (in percentage) among species of all samples from this study and from Genbank, conducted using Tamura-Nei parameter model.**

	<i>C. porosus</i> (Samples)	<i>C. porosus</i> (Samples and genbank)	<i>C. intermedius</i>	<i>C. mindorensis</i>	<i>C. palustris</i>
<i>C. porosus</i> (Samples)	0-0.1				
<i>C. porosus</i> (Samples and genbank)	0.5-1.7	0-1.3			
<i>C. intermedius</i>	9.8-9.9	9.3-10.5	0		
<i>C. mindorensis</i>	8.7-9.9	8.5-10.5	9.6-10.6	0-0.8	
<i>C. palustris</i>	10.0-10.5	9.1-10.6	9.7-10.3	6.9-8.4	0-0.4
<i>Tomistoma schlegelii</i>	20.1-20.3	20.5-21.0	20.4	22.1-22.8	19.8-20.1

Notes:

	Intraspecific
	Interspecific



**Figure 1.0: Phylogenetic tree constructed using COI and cyt gene sequences showing the relationship among *Crocodylus* species with *Tomistoma schlegelii* as outgroup.**

**Bootstrap values next to the node were Maximum Parsimony (MP), Neighbour-joining (NJ) and Maximum likelihood (ML) analysis.**

All nine samples in this study and two samples from Genebank are closer to each other and grouped into one clade namely Clade A, with bootstrap values of 100%, 100% and 99%, for MP, NJ and ML, respectively (Figure 1), thus *Crocodylus porosus* is monophyletic. There is a subclade in Clade A, comprising all samples from Sarawak, which form sister clade relationship with other crocodile samples. Other species of crocodiles are also reciprocally monophyletic.

### Discussions

In this study, CTAB protocol (Doyle and Doyle, 1987) had been successful to extract

total genomic DNA of crocodiles, and the products could be used as templates in PCR. In other studies, researchers used different methods for example standard salting-out protocol (Lapbenjakul et al., 2017), DNeasy blood and tissues kit (Rossi et al., 2020), QIA mg DNA micro kit (QIAGEN, Germany) (Gani et al., 2019), standard proteinase K digestion, phenol chloroform extraction and ethanol precipitation (Li et al., 2017).

In this study, 499 bp of COI gene and 418 bp of Cyt b gene had been successfully sequenced. Li et al., (2017) had reported 1557 bp of COI gene and 1165 bp for Cyt b gene from saltwater crocodiles which unsimilar to this study. Difference of primers have been

used from this study with Li *et al.*, (2017) study. All sequences obtained in this study matched *C. porosus* data in the genebank Accession No.: DQ273698.1 and NC008143.1.

The nucleotide composition in this study showed A and T has higher percentage than C and G, which is similar to Li *et al.* (2007). They reported AT composition is higher than GC. This study recorded nucleotide base in descending order: C > A > T > G but in Li *et al.* (2017) nucleotide composition is as follow: A > C > T > G (Li *et al.*, 2007). According to Borrow (2003), AT content of reptiles and mammals is lower than GC while Meister and Borrow (2005) stated that the same species tend to have same nucleotide composition.

Samples of *C. porosus* had values of genetic distance between zero to 1.7%, which indicated a strong intraspecific variation (Brandley and Baker, 2001). Gani *et al.* (2019) reported genetic divergence values between zero to 0.6 %, among *C. porosus* in Sarawak. Genetic Species Concept by Bradley and Baker, (2010) noted that genetic distance values of more than 2% indicates intraspecific variation, values above 2% to 11% suggest probably a high indicative of conspecific populations whereas if more than 11% indicates species recognition (Bradley and Baker, 2001).

In this study, phylogenetic trees showed that *C. porosus* is monophyletic, similar to those reported by Meganathan *et al.*, (2010). This study suggested that combine analysis of COI and Cyt b gene could resolve phylogenetic relationships within Order Crocodylia, as all species form reciprocally monophyletic groupings. This finding is coherent with current classification using morphological characters.

## Conclusion

Based on combine COI and Cytb gene information, intraspecific variation among samples from Sarawak ranged between zero to 0.1% and with other *C. porosus* samples up to 1.6%. Phylogenetic analysis using combine genes also supported the monophyletic of *C. porosus*. There is no geographical clade of *C. porosus* could be observed during this study. In future, more samples are needed in order to resolve relationships among crocodiles in Sarawak.

## Acknowledgement

The Genetic study on wild crocodile population in Samarahan River Basin is supported by Ministry of Higher Education Malaysia

FRGS/1/2019/WAB13/UNIMAS/02/2.

Authors would like to thank UNIMAS for laboratories and equipment, local guidelines during field sampling.

## References

- Abdul-Gani, M.I.Z. (2014). Population Density, Human-Crocodile Conflict and Genetic Variation among Saltwater Crocodile, *Crocodylus porosus* in Sarawak. (Unpublished Master's thesis). Universiti Malaysia Sarawak, Sarawak, Malaysia.
- Abdul- Gani, M.I.Z. (2019). Historical Information, Population Density and Distribution, Human Crocodile Conflict, Crocodile Habitat, Selected Water Parameters, Genetic Variations among Saltwater Crocodile, *Crocodylus porosus* in Sarawak. [dissertation]. PhD Thesis, Universiti Malaysia Sarawak.
- Abdullah, N. S. S. (2010). Molecular Phylogeny of Saltwater Crocodiles

- (*Crocodylus porosus*) from Sarawak based on Combined Cyt b and 12S rRNA Gene Analysis. (Unpublished Master's thesis). Universiti Malaysia Sarawak, Sarawak, Malaysia.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic Local Alignment Search Tool. *Journal of Molecular Biology*, 215(3), 403-410.
- Bradley, RD & Baker, RJ 2001, 'A test of the genetic species concept: cytochrome-b sequences and mammals', *Journal of Mammalogy*, vol. 82, no. 4.
- Doyle, J. J., & Doyle, J. L. (1987). A Rapid DNA Isolation Procedure for Small Quantities of Fresh Leaf Tissue. *Phytochemical Bulletin*, 19, 11-15.
- Barrow, M. (2003). Relationship between Genome Size, Base Composition and Endopolyploidy in seed plants. Retrieved from doi.org/10.1002/9783527610921.ch8.
- Gani, M., Mohd Kulaimi, N.A., Rosli, N., Sitam, F.T., Ithnin, H., Abdullah, M.L. & Rovie-Ryan, J.J. (2019). DNA forensic case study: species identification from suspected crocodile penis. *Journal of Wildlife and Parks*, 34: In press.
- Grigg, G. C., & Gans, C. (1993). Morphology and Physiology of the Crocodylia. In C. G. Glasby, B.J. Ross & P. L. Beesley (Eds.), *Fauna of Australia Vol 2A: Amphibia and Reptilia* (pp.326-336). Canberra: Australian Government Publishing Service.
- Hassan, R., Ahmad, R., Md Adzhar, M. A. A., Abdul-Gani, M. I. Z., Ayob, A., & Zainudin, R. (2016). Notes on the Wild Tomistoma Populations in Western Sarawak, Malaysian Borneo. *International Journal of Ecology*, 2016, 1-7.
- Hassan, R., Md Adzhar, M. A. A., Abdul-Gani, M. I. Z., & Ahmad, R. (2018). Assessment of Wild Saltwater Crocodile Population in Bako River, Western Sarawak, Malaysian Borneo for Potential Ecotourism Industry. *Malaysian Applied Biology*, 47(1), 131-138.
- Hassan, R., Mohd Azizi, N. F., Md Adzhar, M. A. A., Abdul Gani, M. I. Z., Ahmad, R., & Moi Ung, C. L. (2018). A Taphonomic Study of *Crocodylus porosus* (Crocodylidae) and *Tomistoma schlegelii* (Gavialidae) Remains from Western Sarawak, Malaysian Borneo: Applications for Public Education. *Trends in Undergraduate Research*, 1(1), a23-32.
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7), 1870-1874.
- Lapbenjakul, S., Thapana, W., Twilprawat, P., Muangmai, N., Kanchanaketu, T., & Temsiripong, Y. et al. (2017). High genetic diversity and demographic history of captive Siamese and Saltwater crocodiles suggest the first step toward the establishment of a breeding and reintroduction program in Thailand. *PLOS ONE*, 12(9).
- Li, Y., Wu, X., Ji, X., Yan, P., & Amato, G. (2007). The Complete Mitochondrial Genome of Salt-water Crocodile (*Crocodylus porosus*) and Phylogeny of Crocodylians. *Journal of Genetics and Genomics*, 34(2), 119-128.
- Martin, S. (2008). Global Diversity of Crocodiles (Crocodylia, Reptilia) in Freshwater. *Hydrobiologia*, 595(1), 587-591.
- Mc Aliley, L. R., R. E. Willis, D. A. Ray, P. S. White, C. A. Brochu, and L. D.



- Densmore. 2006. Are crocodiles really monophyletic? Evidence for subdivisions from sequence and morphological data. *Molecular Phylogenetics and Evolution* 39:16-32.
- Md Adzhar, M. A. A., & Hassan, R. (2017). Relationships among *Tomistoma schlegelii* in Malaysia Based on Cyt b - Control Region Gene Analysis. *International Journal of Zoology*, 2017.
- Meister, A. and Barrow, M. (2005). DNA Base composition of plant genomes. In J. Dolezel, J. Greilnuber, and J. Suda (Eds.). *Flow Cytometry with Plant Cells: Analysis of Genes, Chromosomes and Genomes*. New Jersey, USA: John Wiley & Son.
- Meganathan, P., Dubey, B., Batzer, M., Ray, D., & Haque, I. (2010). Molecular phylogenetic analyses of genus *Crocodylus* (Eusuchia, Crocodylia, Crocodylidae) and the taxonomic position of *Crocodylus porosus*. *Molecular Phylogenetics and Evolution*, 57(1).
- Palumbi, S. R., Martin, A., Romano, S., McMillan, W. O., Stice, L., Grabowski, G., University of Hawaii at Manoa., Kewalo Marine Laboratory. (2002). *The simple fool's guide to PCR*. Honolulu, HI: Dept. of Zoology and Kewalo Marine Laboratory, University of Hawaii.
- Roldán, M.I., Heras, S., Patellani, R. and Maltagliati, F., 2009. Analysis of genetic structure of the red shrimp *Aristeus antennatus* from the Western Mediterranean employing two mitochondrial regions. *Genetica*, 136, 1-4.
- Rossi, N., Menchaca-Rodriguez, A., Antelo, R., Wilson, B., McLaren, K., & Mazzotti, F. et al. (2020). High levels of population genetic differentiation in the American crocodile (*Crocodylus acutus*). *PLOS ONE*, 15(7).
- Shoon, A. (2009). Sequencing of Cytochrome b and 12S mtDNA Genes of Saltwater Crocodiles, *Crocodylus porosus* from Sarawak. (Unpublished Final Year Project Report). Universiti Malaysia Sarawak, Sarawak, Malaysia.
- Tisen, O., & Ahmad, R. (2010). *Crocodylus porosus* in Sarawak: Status and Management. Paper presented at the IUCN-SSC crocodile specialist group (CSG) Workshop on Human crocodile conflict, Sabah, Malaysia.