DNA BARCODING AND PHYLOGENETIC RELATIONSHIP OF SHARKS AND RAY FISH OF BANGLADESH

By

Sajid Ahmed
16136016

A thesis submitted to the Department of Department of Mathematics and Natural Sciences in partial fulfillment of the requirements for the degree of
Bachelor of Science in Biotechnology

Department of Mathematics and Natural Sciences
Brac University
December 2019

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Sajid Ahmed
16136016

Approval

This Thesis titled "DNA Barcoding using Cytochrome c oxidase gene Marker of Marine Sharks and Ray fish of Bangladesh" submitted by

1. Sajid Ahmed (16136016)

of Spring,2014 has accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Science in Biotechnology on 5th December 2019.

Examine Committee:

Supervisor 1:
(Member)

Ms. Romana Siddique
Senior Lecturer, Mathematics and Natural Science
Brac University

Supervisor 2:
(Member)

Dr. Md. Sagir Ahmed
Professor, Department of Zoology
University of Dhaka

Program Coordinator:
(Member)

Iftekhar Bin Naser
Assistant Professor, Mathematics and Natural Science
Brac University

Department Head:
(Chair)

Prof A.F.M Yusuf Haider
Chairperson, Mathematics and Natural Science
Brac University

Ethics Statement

The Study is conducted with the samples from the DNA Barcoding Lab, Department of Zoology, Dhaka University. The property rights belong to Dr. Md. Sagir Ahmed (Professor, Department of Zoology, Dhaka University) and was given consent to use it only for this thesis purpose only. The samples and related data are strictly confidential and any further unitarization of these data would require this permission for it first.

Abstract

Objective: The Lack of proper identification of various sharks and ray species has been a far cry among scientists. Cytochrome c oxidase gene-based DNA barcoding is used in various species identification not only because of pinpoint accuracy but how it can be used only a small amount of sample which doesn't have proper morphological identification scope. So a need for a strong database of these sharks and rays identifying marker Cytochrome c oxidase gene is needed to crosscheck, regulate and update for current use and future reference.

Methods: Collection of samples were done from the southern region of Bangladesh in the Bay of Bengal at various points and times. To identify the species a partial portion of it was collected, in general an approximate 650bp sequence of the mitochondrial cytochrome c oxidase subunit I (COI) gene is taken in consideration for this study. Since cytochrome c oxidase gene is easy to multiply in numbers with specified primers, a good number of outcomes are taken into consideration via Sanger-Sequencing method. The raw sequences were then trimmed according to match value and uploaded to NCBI database from where a cross match is done to verify the sequences. Further analysis was done using both offline and online resources of bioinformatics tools.

Results: Our study concluded results that the K2P percentage of intra species of sharks and rays cluster around a close range of 1% to 6.5% in intra species and the range is spread from 11% to a maximum of 64% between species which gives us a genetic gap of 4.5%. For sharks, the GC_1, GC_2, and GC_3 codons were 52.52 ± 0.23, 42.89 ± 0.04, and 20.38 ± 0.31, respectively with an average GC percentage being 38.59 ± 0.12 in the 8 species of sharks. For the rays, the GC_1, GC_2, and GC_3 codons were 53.96 ± 0.36, 43.48 ± 0.19, and 33.59 ± 0.44, respectively with an average GC percentage being 43.67 ± 0.28 among 10 species of rays. Amino acid profiling showed clear similarities among the percentage of amino acid translated from the mitochondrial COI gene of same species and a significant difference between intra and inter species which makes it a good second marker apart from the GC percentage via DNA barcoding method.

Conclusion: This study proved that COI gene-based DNA barcoding is a sound technique for rapid and accurate identification of sharks and rays also amino acid profiling can also be set as possible additional identification marker.

Acknowledgement

The initial gratitude would go to Ms. Romana Siddique who has been a constant guide in my academic life in Brac University and to help me to aim for goals more than academic ranges and beyond to a professional life ahead. She has always been a center for solution, support, encouragement and a calm head to get me through.

I would then like to mention Dr. Md. Sagir Ahmed who has given me the opportunity to work with him and his team in an incredible task of DNA Barcoding of Marine and Freshwater species of Bangladesh. The huge team set up by him and maintained are enriching the DNA database of Bangladesh fish species. I am extremely grateful to be trusted with confidential data of his property right claimed samples of sharks and ray fish and let us work under him and his lab system.

A special thanks to my Co-supervisor Md. Zarif Hossain who has given us proper guideline and training to begin and retrieve results from the beginning of this project work which ultimately resulted as my thesis work.

Also like to acknowledge the chairperson of the Department of Mathematics and Natural Science, Prof A.F.M Yusuf Haider who has been enabler for this department to keep up with the modern times for the need of the future and maintaining a curriculum that satisfies the standards an institute must have.

I would also like to mention my gratitude towards my parents who have been very supportive towards my choices and decisions at times to help me get to this point. Without their encouragement putting effort and time in the line of science would have been very difficult for me.

Table of Contents

Chapter 3 Results

List of Tables

List of Figures

Introduction

The Bay of Bengal which is to the south of Bangladesh stretch a vast amount of area with unique range of biodiversity surrounding it. With many rivers carrying out fresh water towards the sea and having 710km of long coast line with exclusive economic zone of 200 nautical miles, Bangladesh has good “Blue economic” opportunity, a term coined to identify not only marine life harvesting potentiality but more like natural resources, trades etc [1].

Sharks and Rays are observed as very highly misidentified species almost at an identification error rate of 20% for carcharhinid sharks only [3]. Deep sea fishing often catches various sharks and rays while fishing for other fishes commercially but recently shark fins and rays are being considered as a delicacy in different culinary cooking dish to some people of various part of the world. Hence this creating a new window of opportunity for the fishermen to catch certain sharks for their fin and rays for the meat to be sold as an exclusive food item, which is creating a concern of whether these species are overfished or vulnerable species are being captured to the point of extinction. [4,5] Although it is difficult to control the overfishing and wrong capturing of sharks and rays or commercial hunt, proper identification and collective data about habitats of different species and how many of them are present will help us to preserve them from near extinctions of many species and create a buffer period for the biodiversity to recover if overfishing is being damaging it to the point of no recovering. [4,5] Sharks are considered as the top predator of the ocean and at the top of the food chain, but since the global need for various needs for shark fins, meat, liver oil which is full of vitamin A, Shark hide, Rays fin and meat the global capturing and poaching of sharks and rays have risen quite some fold and the similar pattern in visible in Bangladesh [6]. To prevent mass poaching and wrong hunting of endangered species we need proper identification of each species of rays and shark but it is very hard to do a taxonomical Identification without expertise to differentiate closely related species and for that DNA barcoding is used. The preserved region of any given species mitochondrial DNA gene cytochrome C oxidase subunit I gene (cox1 or COI) is taken as standard genetic marker [2]. The overall impact of shark catching is very significant to on 0.96% of the total marine fish landing in Bangladesh but the demand if not in local market but in the global market can give rise to uncontrolled poaching and

hence the need for proper way to identify the trades with DNA barcoding method [7]. During natural calamities like Hurricane, Tsunami or a scarcity of food both sharks and rays being a strong swimmer and predator can migrate to places where they usually don't roam around. This can be another reason for misidentification and the need to have proper use of DNA barcoding data and analysis for proper identification. Another concern of sharks and rays is that illegal poaching often can't be traced down properly as the partial body parts are often recovered like either just the fins or cut flesh at a point where morphological traits cannot be used to differentiate [25]. In this sort of cases DNA barcoding can come as a very useful and effective. Apart from Identification this study will help to know the comparative relation between sharks and rays Kimura-2-parameter distance to see a clear relationship in terms of distance point between the same order of species and interspecies [26].

Even though Bangladesh is a land of huge population and the demand for Protein mainly comes from farmed poultry, cattle and fish from rivers, ponds and artificially cultured, The Bay of Bengal is becoming a huge source for protein with the advancement of fishing boats and material. This might be a good solution to our protein need solution but it creates a great risk of eco system imbalance of the sea and may as well decrease the number of predators such as sharks and rays living freely in less strained food supply areas of the sea. This is why identifying the species living within the area of our sea and having better knowledge about overexploitation and saving the habitat may save these creatures of the ocean from the brink of extinction any time soon.

Evolution among the animals usually take place under a various number of condition and situation. These Evolution can be understood even better if we have time wise mapping of them in such a way that can give us a clear and consistent information of their existence and cross matched with morphological change. Sharks and rays are one of the oldest predators of the sea and has gone though many changes in various period of time. The cross breeding between species and their ancestral roots would be easy to identify by comparing with previous COI gene data base, hence this set of work has many potentials for future work reference.

1.1 Importance of COI in Identification of Sharks and Ray Species:

The revolution to use COI gene as DNA Barcoding marker came down to the feature of COI gene being much conserved region for Animal mitochondrial DNA (mtDNA). Since each cell contains a lot of mitochondria it is easy to detect. The absence of intron in the COI gene region makes it more fixed and unchangeable. Also, the fact that fragment data can be used for DNA Barcoding identification. This location of DNA fragment close to 655bp can be easily located with makes it more suitable for it to be set as a marker [40]. Each fragment is easily be multiplied with a specific primer in PCR method [41]. This COI Region is reserved for each species but the variation ratio is high for separate species of animal. And due to the maternal inheritance there is very little chance of the gene to be manipulated across breeding process. Considering these facts COI gene can be taken as a secure map for identification of species with precision.

1.2 Specific objectives of this study

The specific objectives of this study were:

1. To know the genetic diversity of sharks and rays.
2. To establish Phylogenetic relationship among different groups of sharks and rays.
3. To enrich the Barcode reference library of sharks and rays and evaluating different parameters of COI sequences.

Chapter 2 Material and Methods

2.1 Study area and period

This study was conducted in the regions of southern and Southern east regions of Cox's bazar and Kuakata over the period of various time period of 2017-2018. Collection of samples was done at different range of fishing expedition or collected fish in local markets or from fishermen.

2.2 Procedure of sampling

The collected muscle sample was kept in an ice cooling box while it traveled from collection point to the Advanced DNA Barcoding and Fisheries Laboratory, Department of Zoology, University of Dhaka for further studies. Required amount of DNA was taken out in times of need for experiment and analysis and the sample amount is sufficient for a multiple number of DNA extraction. For long time preservation the samples are put under a nitrogen-based cooling system.

2.3 Tissue Collection and Sample Preparation

1. Each shark and ray specimen was sliced and minced by sharp sterilized blade on the dorsal fin base. About 20 mg (5mm) of fish tissue was taken from the sample on a Petri dish. It was further prepared for DNA extraction. Tissue which was not immediately used was to be preserved into absolute alcohol at -20 °C.

2. Any fat, scale or other unwanted debris were removed from the collected tissue sample. Then it was taken into an Eppendorf tube where it was cut by scissors into pieces as small as possible.

3. The tissue was homogenized gently in 2 volumes (w/v) cold TNES buffer/Lysis buffer. Furthermore, the volume was adjusted to 500 µL.

4. 10 µL of Proteinase K was added to it. It was incubated overnight at 56 °C until the tissue was totally dissolved.

5. This dissolved tissue was used for DNA extraction.

2.4 DNA Extraction

Although the basic steps in the isolation of DNA from tissue are fairly constant, a variety of modifications exist for DNA extraction from aquatic species, including numerous commercially available kits. Oftentimes, the choice of DNA extraction method is dependent on the status of the starting material, and factors such as tissue type and DNA integrity are taken into account. DNA can be damaged by events such as heat exposure, low pH, and nucleases that cause enzymatic degradation, depurination, and hydrolysis. DNA found in processed seafoods may have undergone significant damage,with the result being reduced quality and shorter target sequences than those found in a freshly harvested sample. Therefore, a common challenge in the application of genetic methods to the authentication of commercial fish and seafood products is to obtain DNA of sufficient quality and quantity for downstream analysis.

2.4.1 DNA Extraction

1. Equal volume of Phenol: Chloroform: Isoamylalcohol (25:24:1) was added and mixed thoroughly in an up-and-down fashion for few minutes.
2. The samples were centrifuged at 12000 rpm for 15 minutes.
3. Upper aqueous phase was transferred to a new Eppendorf tube and equal volume of Chloroform: Isoamylalcohol was added. Then it was centrifuged at 12000 rpm for 15 minutes.
4. Upper aqueous phase was transferred to a new Eppendorf tube. Double volume of chilled absolute ethanol (100%) was added to it.
5. The samples were again centrifuged at 12000 rpm for 15 minutes. After this step DNA precipitated as pellet. The supernatant was discarded.
6. To wash the DNA 70% ethanol (v/v) was added. They were again centrifuged at 12000 rpm for 15 minutes. Pellet was procured discarding the supernatant and furthermore air-dried.
7. The pellet was resuspended in Nuclease Free Water (NFW) stored in -20° C for long preservation.

2.5 Gel Electrophoresis and Observation of DNA Bands

1. DNA was mixed with the loading dye bromophenol blue and loaded into 1% (w/v) agarose gel. The gel was prepared using agarose powder and 1X TAE buffer with ethydium bromide being dissolved.

2. The DNA was electrophoresed at 110 V for 30 minutes in 1X TAE buffer. The bands were separated thus and furthermore observed. The marker used for DNA ladder 1 kb plus.

3. The gel was observed in the gel documenter, AlphaImager HP using UV light.

4. Image was retrieved in the computer using AlphaImager HP software. The contrast and the background were manually set to get the best of the views. Whether DNA was extracted in the first hand was decided by comparing the respective bands with a fixed corresponding ladder.

2.6 PCR Amplification

1. The major assay currently used in different fish species identification is based on PCR amplification which requires much less starting material (2 μL of DNA) and exhibits greater versatility and sensitivity [33,34].

2. Amplification of genetic material with PCR was done with Taq polymerase, 2 oligonucleotide primers FishF2 and FishR2, 4 deoxynucleotide triphosphates (dNTPs). The final volume was 25 μL.

3. The PCR involved 3 reaction steps carried out at different temperatures: denaturation (approximately 95°C), annealing (50°C to 60°C), and extension (approximately 72°C). During these 3 steps, the template DNA was first separated into 2 single strands by heat denaturation, then the oligonucleotide primers annealed to complementary sequences on opposing ends of a particular fragment of the template DNA, and lastly Taq polymerase utilized the 4 dNTPs to synthesize multiple copies of the target DNA fragment.

4. 50 cycles of these three steps namely: denaturation, annealing, and extension were performed in the Applied Biosystems PCR machine.

5. For agarose gel electrophoresis 10 μl of DNA products from PCR amplification was loaded in 1.2% agarose gel in 1X TAE buffer. Then after running 30 min, it was observed under UV transilluminator. Furthermore, it was documented with Gel-DOC (AlphaImager HP).

2.7 Purification of PCR Product

1. Same amount of loading buffer was added in the PCR tubes (25 µL of loading buffer in case of 25 µL of PCR being produced). [Table-1]

2. Then the tubes were put into the spinner mix for 15-20s.

3. In the purification tubes the whole amount was transferred.

4. Then the tubes were micro centrifuged for 1 min at 12,000 rpm.

5. The tubes were discarded with the elutes in them.

6. The suckers were taken into the fresh pre-sterilized Eppendorf tubes.

7. 700 µL wash buffer was pipetted into each of the sucker.

8. The tubes were centrifuged for 1 min at 12,000-14,000 rpm.

9. After the liquid being discarded from the columns the tubes were centrifuged for 1 min at 12,000-14,000 rpm for the 2nd time.

10. The used columns were discarded.

11. Finally new columns i.e. capless MC tubes were taken with 40 µL elution buffer in each of them.

12. Then the tubes were centrifuged for 1 min at 12,000 rpm for the final time.

13. The sucker was discarded and the columns were preserved in -20°C. Each of the column contained purified PCR product.

14. The purified PCR products were loaded in the 1.2% (w/v) agarose gel and electrophoresed at 110 V for 30 minutes. The purified PCR products were observed under gel documenter (AlphaImager HP) and thus an electropherogram was produced. [Figure 2.7.1]

Table-1 List of primers used for COI amplification

FishF1	TCAACCAACCACAAAGACATTGGCAC	Forward	**26**
FishR1	TAGACTTCTGGGTGGCCAAAGAATCA	Reverse	**26**
FishF2	TCGACTAATCATAAAGATATCGGCAC	Forward	**26**
FishR2	ACTTCAGGGTGACCGAAGAATCAGAA	Reverse	**26**

Table-2 Composition of the Master Mix

Components	Volume
1.Taq polymerase	**12.5**
2. Forward primer	**1**
3. Reverse primer	**1**
4. Extracted DNA	**2**
5. Nuclease free water	**8.5**
Total	**= 25 µl**

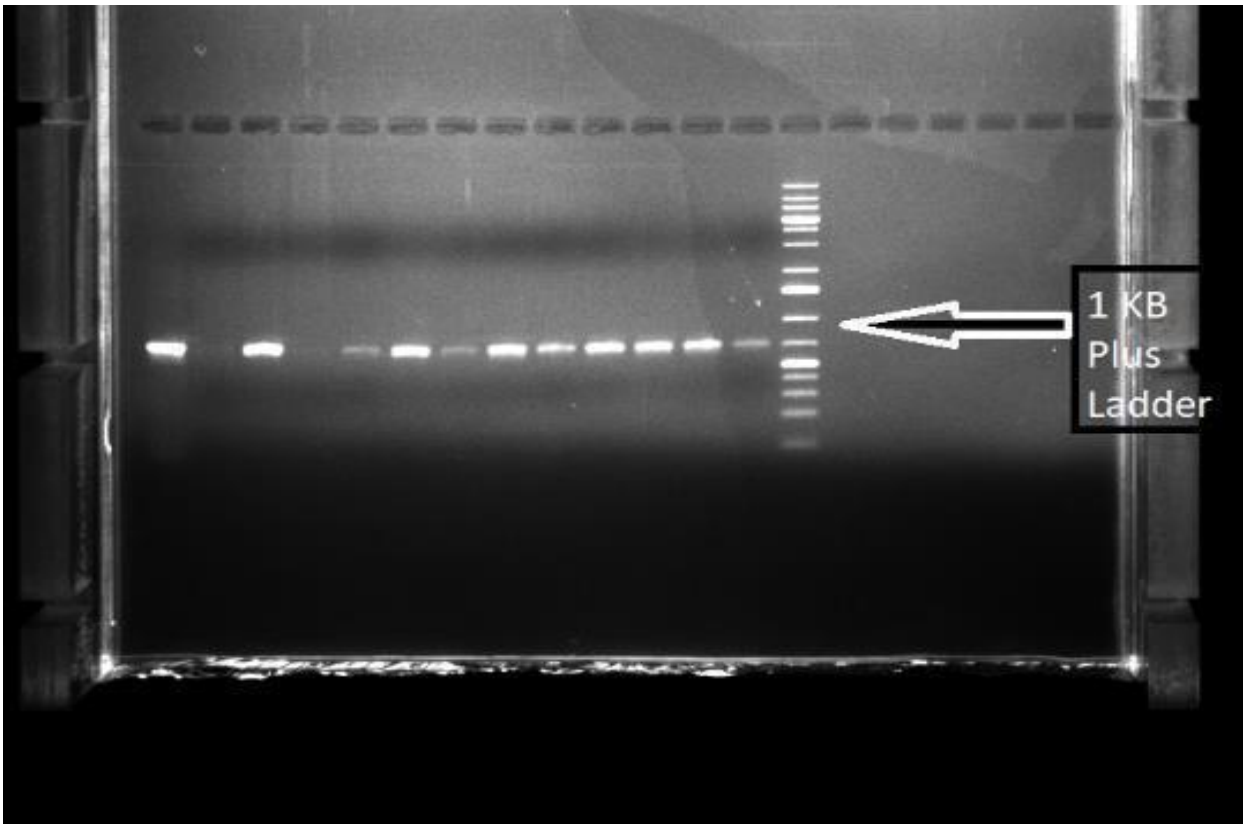

Figure 2.7.1: Electropherogram obtained after the elecctrophoresis of PCR product.

2.8 DNA Quantification and Sequencing

1. Quantitation of DNA was estimated in three different absorbances and the average was taken into consideration. Nano drop spectrophotometer was used for this purpose.
2. After that, the purified PCR products were sent to the First BASE laboratories, Malaysia for sequencing.

2.9 NCBI bankIT sequence submission

1. After Sanger sequencing from the Malaysia a chromatographic data is received which looks like colorful barcode hence the name DNA barcoding.
2. Each sequence is needed to be trimmed on the upstream portion and downstream portion around 20-30 bp on average depending on where the quality match score gets better.
3. Trimmed sequence is converted into FASTA format of file and with proper information charts and primer list the sequence is processed for upload in Genebank with bankIT tool.
4. ExPASy tool for DNA sequence translation were used as the translated amino acid sequence is required to upload the sequence.
5. Each Batch of sequence upload gets crosschecked from the receiving end and an accession number is given for each sequence of species.

2.10 Bioinformatics Analysis

1. After the filtering of just shark and ray sequences from the other species batch lot. Each sequence was rearranged according to its Order and pasted in a FASTA format file for the MEGA-X
2. Grouping was done according to the species name for species wise K2P distance, GC content and phylogenetic tree construction, another set of grouping was done according to the Order.
3. The results were exported to an excel file for further inspection and chart creation for visual representation.
4. For the analysis purpose more sequence FASTA of similar species were downloaded from NCBI website database as secondary reference data.

2.11 Process of amino acid profiling

After making a FASTA file containing the similar species, the file was opened from MEGA-X and aligned and analyzed with MUSCLE algorithm. The aligned file was saved in a session and translated to amino acid with help of pre-built modeling option which lets you convert Nucleotide codons in amino acid in percentage form. From that data set species wise and amino acid wise separate bar chart and cluster chart was built where to give each graph better visual effect each value was multiplied with 10. Separate rows and column were built for visualizing specific amino acid present at different percentage in various shark and ray species. [27]

2.11.1 Amino acid analysis

The Nucleotide sequence were aligned first using the muscle codon algorithm of the MEGA X software where the option to convert the codons into amino acid expression was selected. The aligned amino acid were retrieved in excel file in percentage ratio of amino acid for each sequence. The graphical representation was done using the excel chart to show a clear pattern of amino acid range and to find anomaly among the sequences.

2.11.2 Amino acid analysis using Swiss protein modeling and ExPASy Translate

If there Translation of the nucleotide sequence is done in the ExPASy translate website and the result is put into the Swiss protein modeling to see if the overall position of the amino acid distribution, proline, glycine and pre-proline distribution are same compare to an established reference sequence and the sequence that the species might be related to. This helps to pinpoint what is sequence can be related to on a molecular level.

Chapter 3 Results

Sharks and rays sequences were categorized according to species and order and 8 species of sharks and 10 species of rays in total 18 species were taken in consideration for this analysis. The total number of raw sequences used were 44. Furthermore, each sequence was used in BLAST to find similar sequence around the world and several sequence for each species was downloaded as secondary reference data.

Table-3 List of shark species with their IUCN red list state and accession number

	Scientific Name	English Name	IUCN Red List Status	Accession Number
1.	*Chiloscyllium burmensis* (Order: Orectolobifomes)	Burmese Bamboo Shark	Data Deficient (DD)	1.MH429291.1 2.MH429293.1 3. MN083134
2.	*Carcharhinus amboinensis* (Order: Carcharhiniformes)	Pigeye Shark	Near Threatened (NT)	4.MH230957.1
3.	*Carcharhinus sorrah* (Order: Carcharhiniformes)	Spot-tail Shark	Near Threatened (NT)	5. MH429287.1 6. MH429296.1
4.	*Carcharhinus leucas* (Order: Carcharhiniformes)	Bull Shark	Near Threatened (NT)	7. MH230955.1
5.	*Scoliodon laticaudus* (Order: Carcharhiniformes)	Spadenose Shark	Near Threatened (NT)	8. MH230956.1 9. MH429292.1 10.MH087057.1 11.MH087056.1
6.	*Galeocerdo cuvier* (Order: Carcharhiniformes)	Tiger Shark	Near Threatened (NT)	12. MN013428 13.MH429290.1
7.	*Rhizoprionodon oligolinx* (Order: Carcharhiniformes)	Grey Sharpnose Shark	Least Concern (LC)	14.MH311285.1 15.MH429295.1 16.MH311281.1 17.MH311279.1
8.	*Sphyrna lewini* (Order: Carcharhiniformes)	Scalloped Hammerhead Shark	Endangered (EN)	18.MH230949 119.MH429288 120.MH429289.1

Table-4 List of ray species with their IUCN red list state and accession number

	Scientific Name	English Name	IUCN Red List Status	Accession Number
1.	*Narcine maculata* (Order: Torpediniformes)	BlackSpotted Numbfish	Data Deficient (DD)	21.MN083137
2.	*Narcine brunnea* (Order: Torpediniformes)	Brown Numbfish	Not Evaluated (NE)	22.MH882465.1 23. MH882464.1 24. MH429319.1 25. MK792776.1 26. MN083105
3.	*Himantura walga* (Order: Myliobatiformes)	Scaly Whipray	Near Threatened (NT)	27. MH429304.1 28. MH429305.1 29. MH429306.1 30. MH429310.1 31. MH882466.1 32. MN013425
4.	*Himantura undulata* (Order: Myliobatiformes)	Honeycomb Whipray	Vulnerable (VU)	33. MN013427
5.	*Himantura gerrardi* (Order: Myliobatiformes)	Whitespotted Whipray	Vulnerable (VU)	34. MH230945.1
6.	*Neotrygon kuhlii* (Order: Myliobatiformes)	Blue-spotted Stingray	Data Deficient (DD)	35. MN013424 36. MN013426
7.	Pateobatis uarnacoides (Order: Myliobatiformes)	Whitenose Whipray	Vulnerable (VU)	37. MH230950.1 38. MH230953.1 39. MH230951.1
8.	Pateobatis jenkinsii (Order: Myliobatiformes)	Jekins Whipray	Vulnerable (VU)	40. MH230946.1
9.	*Mobula mobular* (Order: Myliobatiformes)	Giant Devil Ray	Endangered (EN)	41. MH230952.1
10.	*Gymnura poecilura* (Order: Myliobatiformes)	Longtail Butterfly Ray	Near Threatened (NT)	42. MH230947.1 43. MH429307.1 44. MH429308.1

3.1 Morphological Characteristics of the Orders and Family of Shark Samples

Order: **Orectolobifomes**

Sharks which are also known carpet fish have specialized nostril and mouth. Behind the origin of pectoral, the fourth and fifth Gill-opening is visible. Nabial barbels, labial furrows, are at the corner of the mouth region. There are two dorsal fins without the spine. Both anal fin and spiracles are visible. This order of sharks contains both extremely small size and large-scale sharks which roam around deep water. They can be viviparous and oviparous [8,9].

Family: **Hemiscylliidae**

The family is also known as long tailed carpet sharks which has slender and extended body with small mouth which is ventral. It has lower lip- grooves covering the chin. There are large spiracles underneath the eye region and the nostrils have short pointed barbel. There are no lateral dermal ridges on sides of the body [9,10].

Species: ***Chiloscyllium burmensis***

This species is shown to be data deficient in IUCN red list of threatened species list [11]. The English name for it is Grey Bamboo Shark and the local name is Muichya Hangar. This is a bottom dwelling species which feeds on the bottom living invertebrates and small fishes. Usually uncommon in Bangladesh and have less importance in terms of fishery [12].

Order: **Carcharhiniformes**

Sharks which are known as ground sharks usually have similar nostril of ordinary shark without the nasoral groove, perinasal folds or barbels. They have developed labial furrow that the distant end of the jaw. The two dorsal fins visible with the anal fin. The eye sockets are circular in shape and the eyelids are prominent. They happen to be a strong swimmer with five gill slits and often comes close to the river mouth from their usual deep-water roaming region [13].

Family: **Carcharhinidae**

The family of shark is also known as Requiem sharks with elongated body and circular eyes. One of the key features is that the set of teeth is bladelike sharp with one cusp. Other features are having no spiracles, nictitating membrane is present and the fifth gill opening is behind the pectoral fin [9,14].

Species: ***Carcharhinus amboinensis***

A massive sized shark with greyish color skin with white beneath skin color. Roams around surfing level water close to inshore. Key features would be broad and blunt snout, small eyes and large, triangular, saw-edged upper teeth [15]. According to IUCN red list of the threatened species list it is listed as near threatened (NT) [16].

Species: ***Carcharhinus sorrah***

The English name for this species is Spot-tail Shark and the local name is Kamot,Hangor [9]. This species is shown as neat threatened (NT) in IUCN red list of the threatened species list [17]. A small to medium sized shark with a long, rounded snout, large circular eyes, and oblique-cusped, serrated teeth, about 25 teeth is in each jaw. It is colored greyish from above and whitish below. Behind the pelvic fin a dark band mark is visible [9].

Species: ***Carcharhinus leucas***

Usually roams around shallow water near the bays, and rivers. Since it prefers to be in low salinity water it gets the greatest number of encounters with human. Viviparous in nature, attains sexual maturity in 10-15 years. It has massive black snout, triangular shaped upper teeth and small eyes which can be taken as key indicator of this species. The English name of this shark is Bull Shark.

Species: ***Scoliodon laticaudus***

There are three different names for *Scoliodon laticaudus:* Spanose dog, Yellow Dogfish, and Dog shark. It has spade like long snout, slender weak body and eyes are closer to mouth than to the snout. The base of the anal fin is 2-3 times longer than the second dorsal fin [19]. The local name for it is Churi Hangor, Thutte Hangor. On average the size of this shark is 75cm.

Species: ***Galeocerdo cuvier***

The English name of this shark is Tiger Shark and the local name is Bagha Haangor. It has a short and rounded snout and a large head. The slender body with upper furrow prolongated till the eye, large spiracles and set of teeth inclined inwards. The most prominent feature is black strips which makes it look tiger like [20].

Species: ***Rhizoprionodon oligolinx***

The English name of the shark is Grey Sharpnose shark. It is listed as least concerned species in IUCN red list of the threatened species list. The anal fin is in front of the second dorsal fin with short pectoral fin. It roams around both close to shore and off shore, viviparous in nature and identified as harmless to human. The maximum length is of the female of 70cm.

Family: **Sphyrnidae**

Also known as hammerhead shark has an elongated body and moderately slender. Head flattened in front and expanded sideways to resemble a hammer. Mouth inferior, labial furrows vestigial or absent. Teeth blade-like with a single cusp. Eyes at extremities of expansions. Well-developed nictitating lower eyelids. Two dorsal fins, the first high and pointed. Caudal fin strongly asymmetrical, with a well-marked subterminal notch and a small well-defined lower lobe. Precaudal pits present. These are very ferocious sharks.

Species: ***Sphyrna lewini***

This shark has one of the distinctives feature of having head shaped like a hammer hence the English name Scalloped Hammerhead Shark. The local name for it is Gol Kaunna, Juliamagar. Its anterior contour slightly convex with a shallow but distinct indentation at midline. Two more indentations flank the central indentation, giving it a “scalloped” appearance. Mouth broadly arched. Teeth oblique, with a notched outer edge. First dorsal fin high, very dark, falcate its height equal to the length of the pectoral. Second dorsal fin small, low, with a long inner margin and a straight or shallow concave posterior margin, its base noticeably smaller than anal fin base. Pelvic fins low, not falcate with a nearly straight hind margin. Anal fin base noticeably larger than the second dorsal fin base. According to IUCN red list of the threatened species list it is listed as Endangered (EN) [21].

3.2 Morphological Characteristics of the Orders and Family of Ray Samples

Order: **Torpediniformes**

These are rays with electric organs hence the English name Electric Rays. These organs are derived from bronchial muscle in the head region. Their skin is soft and loose and have small eyes. The pectoral disc is large, round and shovel shaped [9].

Family: **Narcinidae**

Thesis family of electric rays are also known as numb fish which is a type of Batoid fish with large oval and shovel shaped pectoral disc. They have shark like tail and jaws are long and broad with teeth remaining exposed when their mouth is closed [22].

Species: ***Narcine maculata***

This species is also known as Darkfinned Numbfish. It is very small in size, on average smaller than 35 cm. This species is considered as Chinese medicine [23]. Mostly roams around continental water.

Species: ***Narcine brunnea***

The English name for this species is Brown Numbfish and roams mainly in continental water. Very small is size and under 22 cm. The local name is Chor Chata, Badami Machh. Electic organs are situated on each side within the disc body.

Order: **Myliobatiformes**

The disc is strongly depressed and varies from oval longitudinally to much broader than long; the tail is well marked off from the body sector, very short to long and whip like, and equipped with a poisonous spine in some species; the pectoral rays are either continuous along the side of the head or separate from the head

Family: **Dasyatidae**

Disc are large in size, circular and have a whip like tail. The body usually consists of denticles, thorns and tubercles. They don’t consist of any dorsal and caudal fins.

Species: ***Himantura walga***

This species is known as Dwaft Whipray or Scaly Whipray. Some the prominent features are that it has smaller eyes in adult form than in young. Teeth are small and dental lamina undulated. It has a whip like tail, greyish brown in color and the pelvic fin is small. The maximum length so far recorded is 45 cm. The local name of it is Hauspata, Sankush.

Species: ***Himantura*** undulata

Venomous rays with a leopard like print on the disc. This creature is oviparous in nature. The size range of it is around. 85-90 cm. The local name for it is Haus, Sankaus, Chitra Haus. Usually considered as a bottom dweller and get caught in the fishermen net by mistake, there is no local demand for it even though in china it is used as medicine ingredient.

Species: ***Himantura gerrardi***

The English name is Whitespotted Whipray which because the tail is spotted with black and white stripes. Vulnerable according to the IUCN red list and is marked as harmless to human. Oviparous in nature with embryos feeding on the yolk initially. Bottom dweller with an average size of 64 cm.

Species: ***Neotrygon kuhlii***

It is called blue spotted stingray. The local name for it is Sapla Pata, Haush, Sakush. The disc is kite shaped with rounded snout. The spiracles are larger than the eye size. The tail is longer than the disc itself and blueish spot it hence the naming. The jaws are small with blunt cuspidate teeth. Usually inhabits in sandy area among coral reefs and moves in deep water as well.

Species: ***Pateobatis uarnacoides***

It is called the whitenose whipray which are usually harmless to human. It exhibits an oviparous nature and on average size of 75 cm. It is popular for its meat and skin as it has high value.

Species: ***Pateobatis jenkinsii***

The local name for it is Golden Crown. Exibits ovoviparity with the maximum size of 104cm. It is found inshore or in sandy area of the sea.

Family: **Mobulidae**

Large pelagic filter feeding rays with a large wide disc and 5 gill slits on the underside. The pectoral fins extend forward forming 'horns' in front of the head that can be uncurled to scoop food into the wide mouth.

Species: ***Mobula mobular***

This is probably the largest rays of all with maximum size of 520cm recorded. An epipelagic species found over continental shelves and near oceanic islands.

Order: **Myliobatiformes**

The primary locomotor organs in most rays, order Myliobatiformes, are the greatly enlarged pectorals which form a wide lateral expansion of the body. Waves of undulation in a vertical plane are passed backward along the mobile fin margins

Family: **Gymnuridae**

The family Gymnuridae contains two genera and about 12 species. Their pectoral fins form a disc that is much broader than it is long, forming wings that give the family its common name, butterfly rays.

Species: ***Gymnura poecilura***

Body compressed and thin, pectoral fin expanded–rhomboid in shape. Tail slender and short; less than disc length; from cloaca to tail tip nearly as long as snout–vent length. The English name for it is Longtail butterfly.

3.3 GC Content variation between shark and ray species

GC percentage is considered as an indicator for identification of both species and order. The FASTA format of the shark sequences were compiled in one FASTA file and put in the MEGA X which is a molecular evolutionary genetic analysis machine which can then analyze and align all the FASTA format files with Muscle algorithm.

3.3.1 GC content pattern of sharks

The average GC percentage of 8 shark species was 38.59 ± 0.12. In our study we can see that the GC_1, GC_2, and GC_3 were 52.52 ± 0.23, 42.89 ± 0.04, and 20.38 ± 0.31 respectively. It clearly showed a trend where $GC_1 > GC_2 > GC_3$. The pattern of standard error of mean (SEM) at different codon positions was also observed and it was 3rd>1st>2nd. SEM for 3rd, 1st and 2nd codon positions were 0.31, 0.23 and 0.04. This sort of similar pattern is followed in the study carried out by other papers [24].

3.3.2 GC content pattern of rays

The average GC percentage of 10 ray species was 43.67 ± 0.28. In our study we can see that the codons GC_1, GC_2, and GC_3 were 53.96 ± 0.36, 43.48 ± 0.19, and 33.59 ± 0.44 respectively. It clearly showed a trend where $GC_1 > GC_2 > GC_3$. (Figure: 3.3.2.1) The pattern of standard error of mean (SEM) at different codon positions was also observed and it was 3rd>1st>2nd. SEM for 3rd, 1st and 2nd codon positions were 0.44, 0.36 and 0.19 [24].

3.3.3 GC content pattern of sharks and rays combined

The combine GC values of shark and ray where overall GC content is 41.41±0.187. The value trend of codons GC_1, GC_2, and GC_3 the usual $GC_1 > GC_2 > GC_3$ in combined sharks and rays and the values were for sharks and rays combined are 53.32±0.17, 43.22±0.082, and 27.72±0.41 respectively. The pattern of standard error of mean (SEM) at different codon positions was also observed and it was 3rd>1st>2nd.

Table-5 List of shark species with their GC percentage on average, 1st, 2nd and 3rd codon

	Scientific Name	Overall GC	1st Codon	2nd Codon	3rd Codon
1.	***Chiloscyllium burmensis***	**37.269**	**47.668**	**42.449**	**21.67**
2.	***Carcharhinus amboinensis***	**38.318**	**52.818**	**42.87**	**19.29**
3.	***Carcharhinus sorrah***	**38.214**	**52.738**	**42.812**	**19.141**
4.	***Carcharhinus leucas***	**38.471**	**53.708**	**42.748**	**18.995**
5.	***Scoliodon laticaudus***	**39.445**	**53.233**	**43.486**	**21.594**
6.	***Galeocerdo cuvier***	**38.075**	**53.882**	**42.595**	**17.747**
7.	***Rhizoprionodon oligolinx***	**38.269**	**53.452**	**42.890**	**18.537**
8.	***Sphyrna lewini***	**40.679**	**52.694**	**43.326**	**26.100**
	Average	**38.59**	**52.52**	**42.89**	**20.38**

Table-6 List of ray species with their GC percentage on average, 1st, 2nd and 3rd codon

	Scientific Name	Overall	1st Codon	2nd Codon	3rd Codon
1.	***Narcine maculata***	**41.003**	**49.908**	**43.913**	**29.138**
2.	***Narcine brunnea***	**38.166**	**45.333**	**41.621**	**27.716**
3.	***Himantura walga***	**44.701**	**56.538**	**42.722**	**34.87**
4.	***Himantura undulata***	**43.64**	**55.241**	**42.85**	**32.87**
5.	***Himantura gerrardi***	**45.812**	**57.164**	**43.133**	**37.152**
6.	***Neotrygon kuhlii***	**43.517**	**55.071**	**42.664**	**32.809**
7.	***Pateobatis uarnacoides***	**45.016**	**57.39**	**43.022**	**34.637**
8.	***Pateobatis jenkinsii***	**45.101**	**56.330**	**42.65**	**36.31**
9.	***Mobula mobular***	**40.909**	**51.998**	**43.175**	**27.535**
10.	***Gymnura poecilura***	**48.893**	**54.717**	**49.082**	**42.88**
	Average	**43.67**	**53.96**	**43.48**	**33.59**

Table-7 Sharks and rays total GC content compared with separate GC contents.

	Overall	1st Codon	2nd Codon	3rd Codon
Rays	43.67±0.28	53.96±0.36	43.48±0.19	33.59±0.44
Sharks	38.59±0.12	52.52±0.23	42.89±0.04	20.38±0.31
Sharks and Rays	41.41±0.187	53.32±0.170	43.22±0.082	27.72±0.41

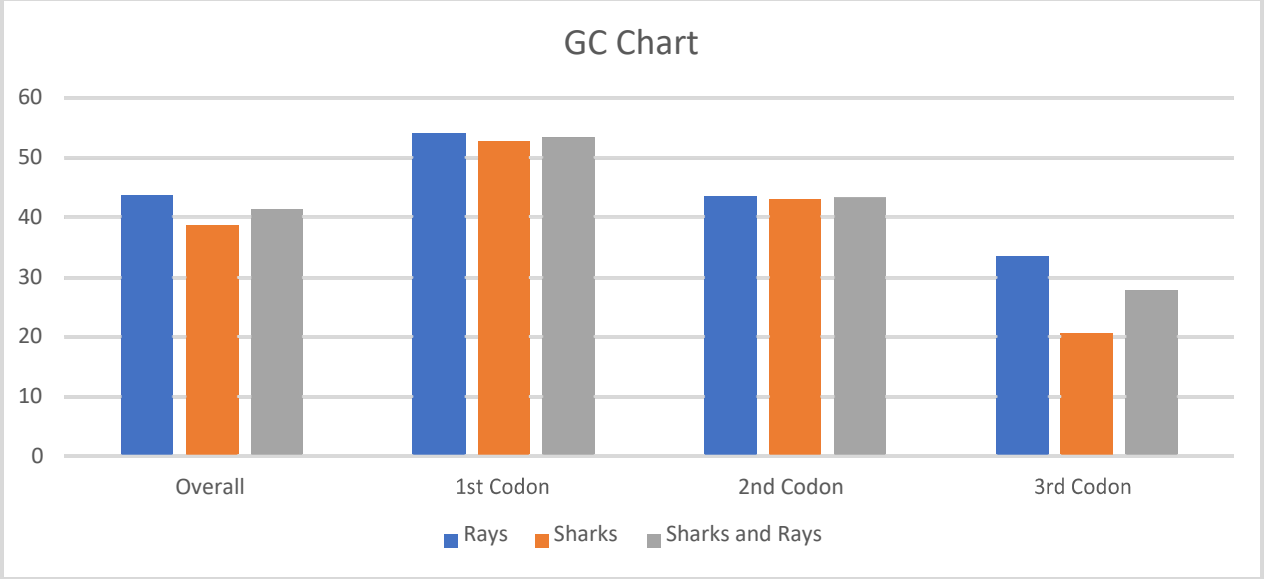

Figure 3.3.2.1: GC overall, 1st codon, 2nd codon and 3rd codon percentage of sharks, rays and both combined.

3.3.4 Whisker box plot of different species of shark GC content

The plot of whisker box plot was done by analyzing the FASTA sequence and taking the average value of GC value across each species. The maximum and minimum value of all GC, GC_1, GC_2 and GC_3 were determined using excel application pre-set formula. The quartile values of first and third were determined with same data set. The whisker box plot represents the lowest value as the

minus downward whisker, followed by the first quartile, median and third quartile as the top of the minus bar to the end of the green box mark and top of the purple box mark. The maximum value is represented as the positive whisker head.

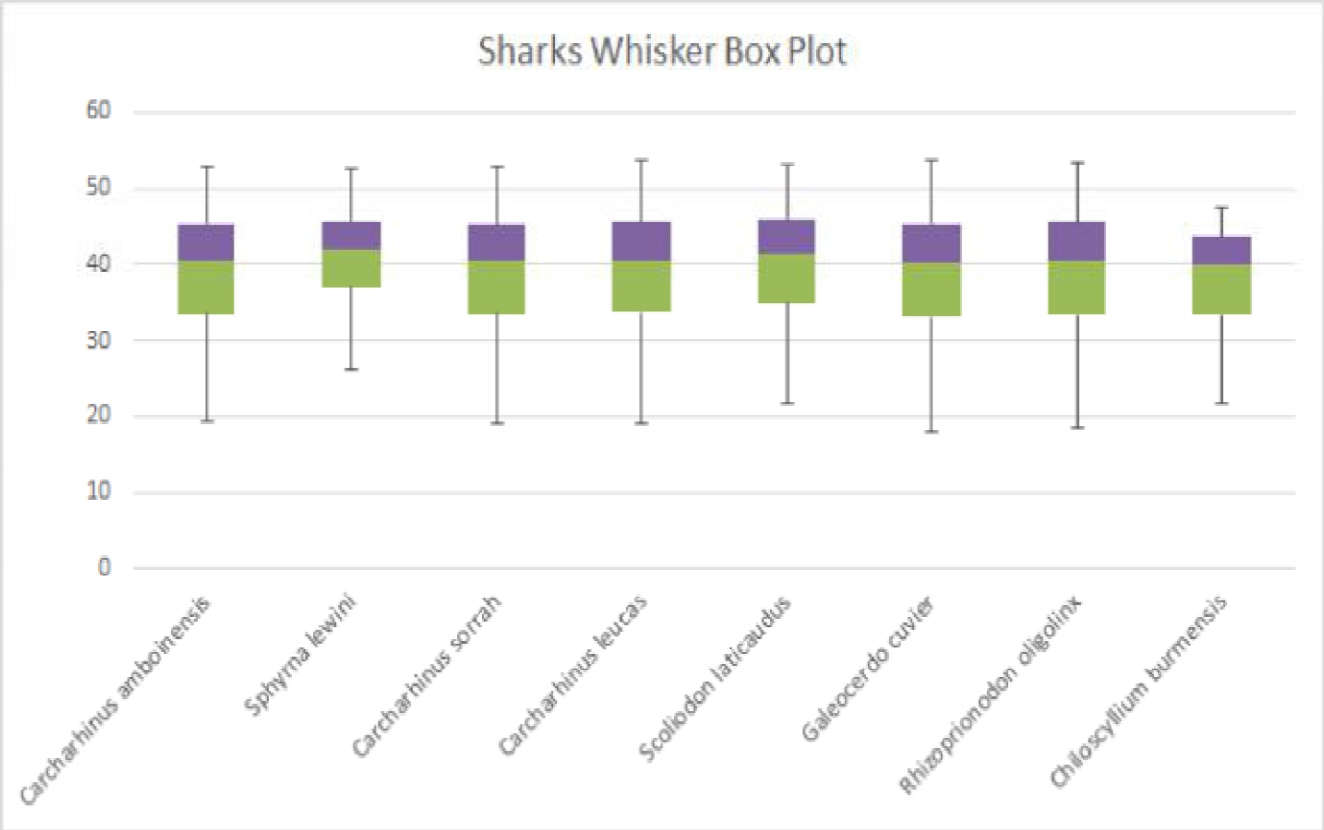

Figure 3.3.4.1: Whisker box plot of different species of shark GC content

3.3.5 Whisker box plot of different species of rays GC content

The plot of whisker box plot was done by analyzing the FASTA sequence and taking the average value of GC value across each species. The maximum and minimum value of all GC, GC_1, GC_2 and GC_3 were determined using excel application pre-set formula. The quartile values of first and third were determined with same data set. The whisker box plot represents the lowest value as the minus downward whisker, followed by the first quartile, median and third quartile as the top of the minus bar to the end of the green box mark and top of the purple box mark. The maximum value is represented as the positive whisker head.

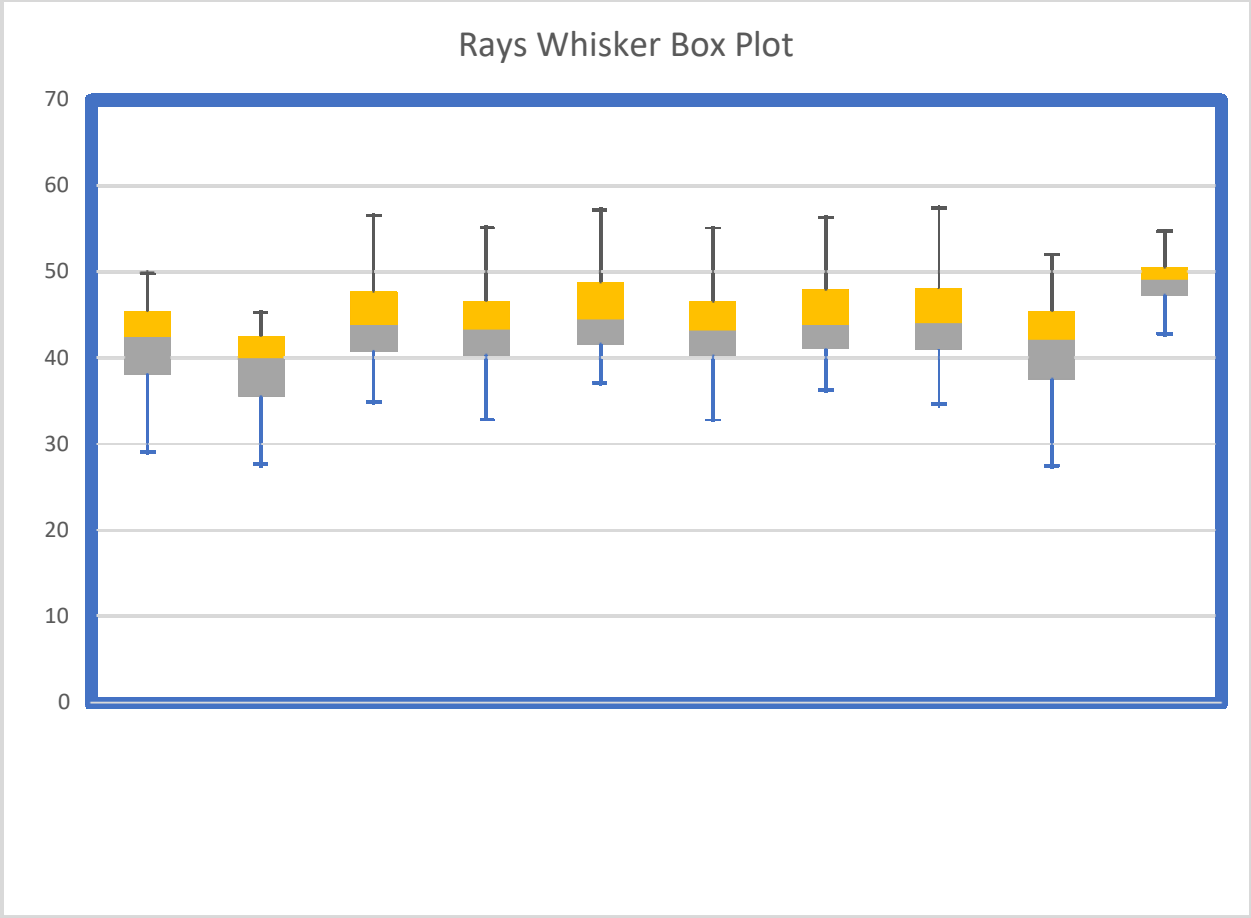

Figure 3.3.5.1: Whisker box plot of different species of rays GC content.

3.4 K2P Distance of Intra and Inter species and Frequency Percentage

For calculating the distance between the inter and intra Order K2P distance some more sequence of similar species was taken from gene bank hit list. For the purpose of comparing data set and creating a more accurate data range of the distance a number of reference FASTA sequence were taken in consideration using MEGA-X and the outcome of the average Intra species distance K2P% was 21% ±0.0008 and average Inter species distance was 29.5 ±0.0005 where the average frequency percentage were 31.16 % for Intra and 5.06% which is close by to the frequency percentage of the mean value. In this set of data, the median value for intra and inter species were 2% and 25.8%. The Graphical presentation of side by side visual representation shows a genetic gap between Intra and Inter species K2P distance it is around the value of 4.5% [Figure 3.4.1, Table-8] , which indicated the clear sign that DNA barcoding works to differentiate between and within species. Also, the cluster set of Intra species K2P distance being close by and the inter species K2P distance shows good variation to further prove that the COI gene-based DNA barcoding works.

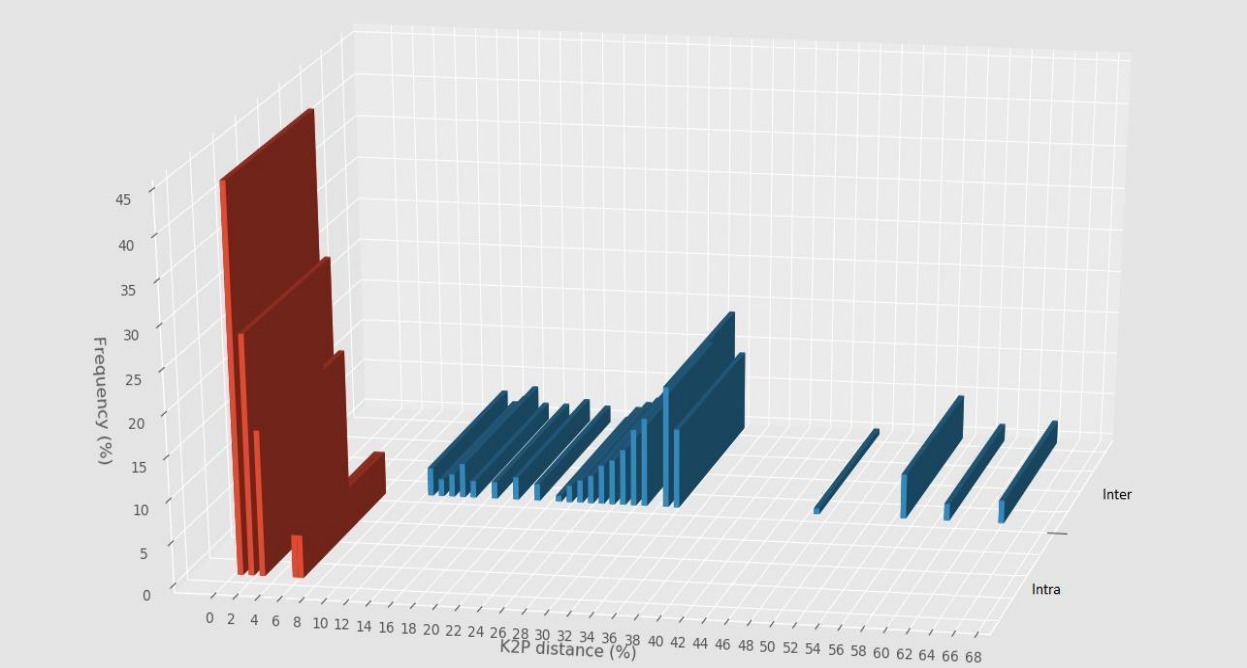

Figure 3.4.1: Frequency distribution of K2P distances percentage of intra species and inter species where genetic gap is 4.5%.

Table-8 Minimum, Maximum, Mean, and Median K2P distance percentage

	Mean	Min	Max	Median
Intra Species	2.1±0.0008	1.1	6.5	2.0
Inter Species	29.5±0.0005	11	38.0	25.8

	Mean	Min	Max	Median
Intra Family	5.9±0.016	0	17.8	3.70
Inter Family	28.7±0.002	14	37.6	29.3

	Mean	Min	Max	Median
Intra Order	7.5±0.015	0	19.0	6.5
Inter Order	30.2±0.002	25.9	37.6	30.1

Table-9 Minimum, Maximum, Mean, and Median K2P distance percentage (with additional NCBI related sequence added)

	Mean	Min	Max	Median
Intra Species	1.31±0.0007	0.00	5.0	1.0
Inter Species	30.4±0.0009	11.0	63.0	29.4

	Mean	Min	Max	Median
Intra Genus	6.06±0.014	0.36	14.2	3.6
Inter Genus	30.5±0.008	12.7	42.6	29.5

	Mean	Min	Max	Median
Intra Family	9.5±0.018	0.47	46.4	5.8
Inter Family	30.6±0.003	13.9	44.8	29.4

	Mean	Min	Max	Median
Intra Order	18.7±0.04	10.3	46.4	14.1
Inter Order	36.2±0.012	26.1	44.9	34.7

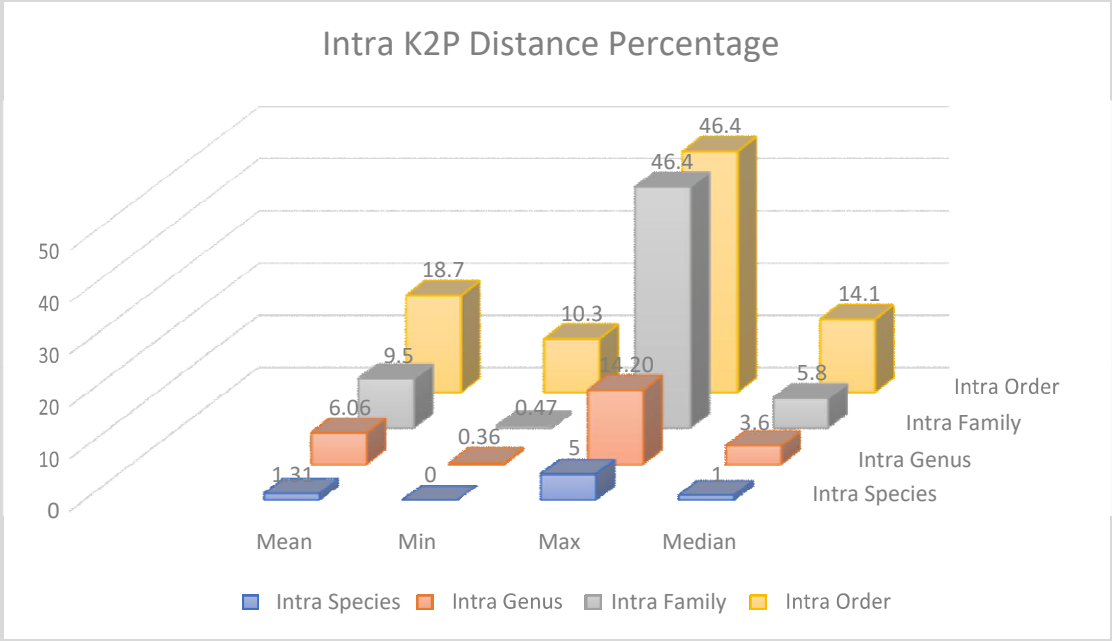

Figure 3.4.2: K2P distances percentage of intra species, genus, family and order.

3.5 Amino acid profiling and a frequency check of COI Gene

Although GC content checking and percentage measurement is considered as the main root of DNA barcoding and identification method but there is often the problem of frameshift or Nucleotide replacement or misidentification of few nucleotide which can give different set of codon variation numbers at times. Even the algorithm to be used can play a role in the points given as GC number as there is different set of point giving criteria gaps, mismatch, and match. With that in mind we have chosen another parameter to make it as a marker. This why amino acid profiling was done on each species to find any pattern of resemblance [27,42].

Amino Acid Composition of *Chiloscyllium burmensis*:

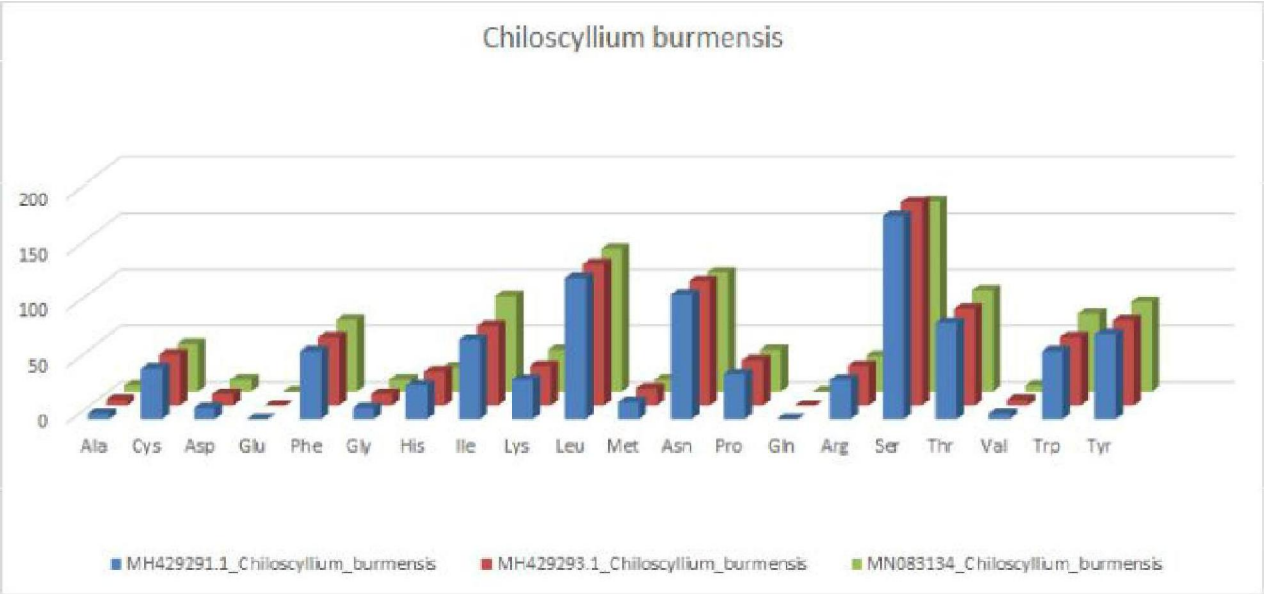

Figure 3.5.1: 20 amino acids percentage (Multiplied by 10) of *Chiloscyllium burmensis*

Amino Acid Composition of *Sphyrna lewini*:

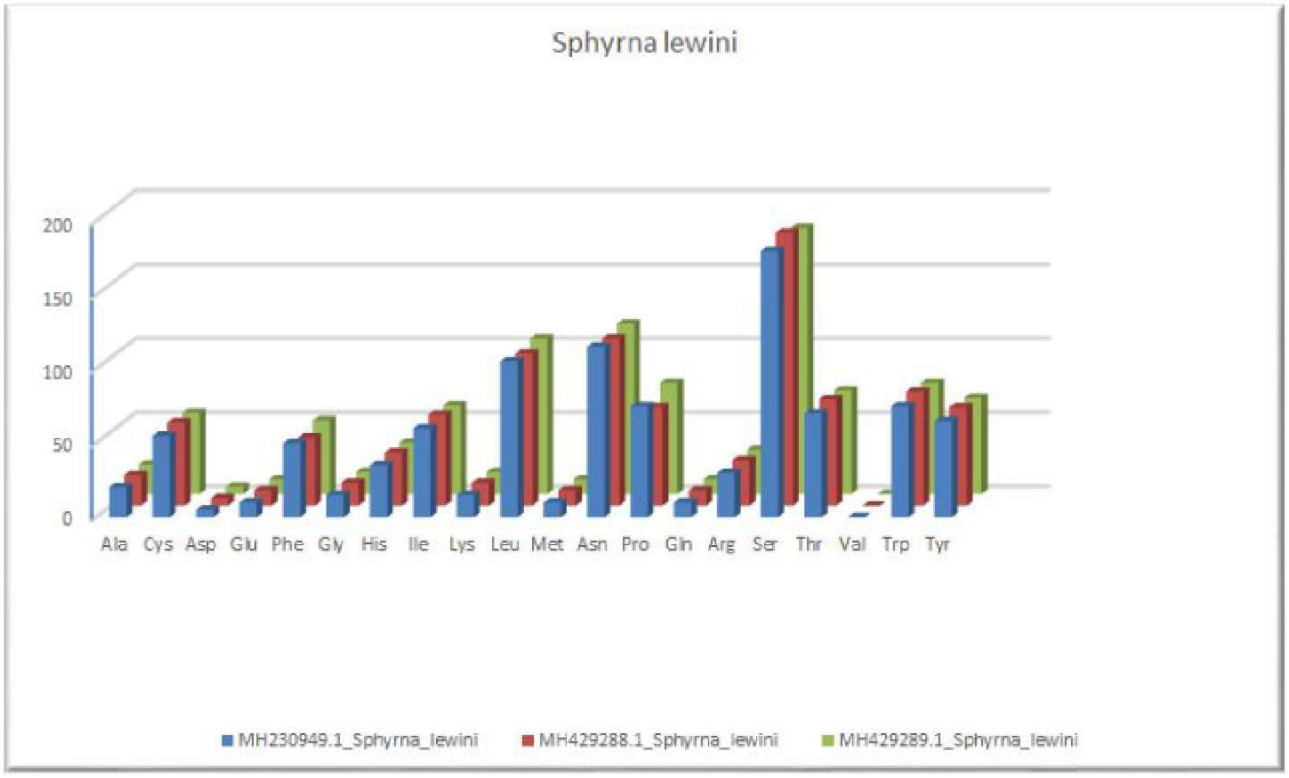

Figure 3.5.2: 20 amino acids percentage (Multiplied by 10) of *Sphyrna lewini*

Amino Acid Composition of *Carcharhinus sorrah*

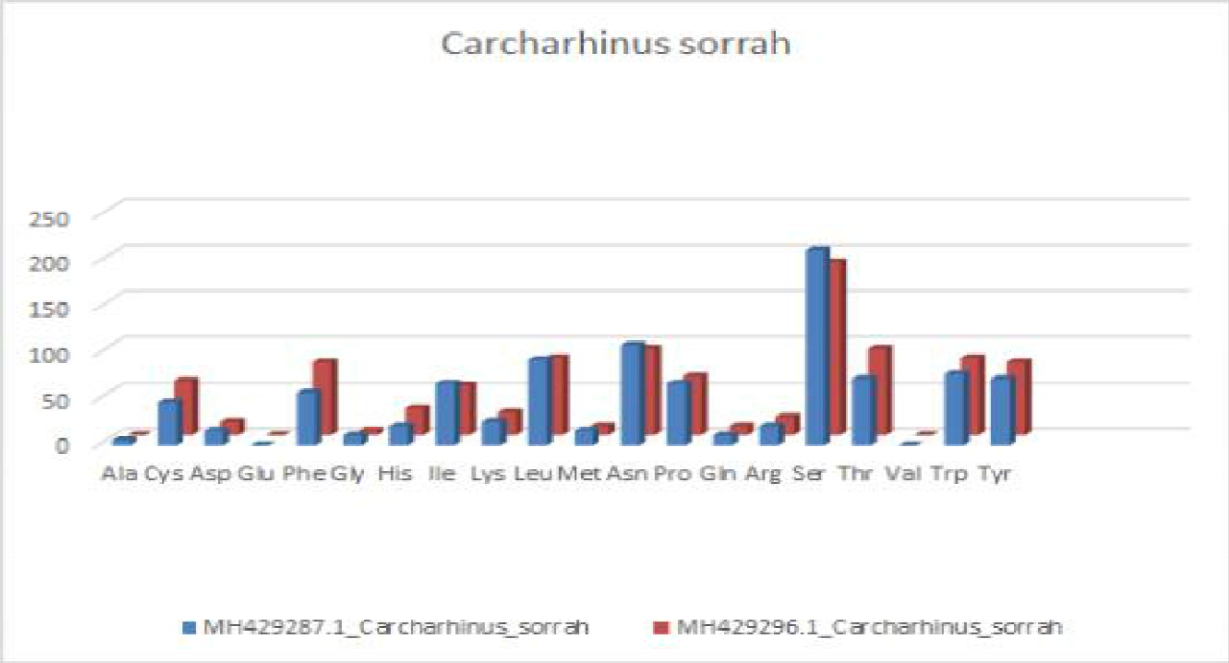

Figure 3.5.3: 20 amino acids percentage (Multiplied by 10) of *Carcharhinus sorrah*

Amino Acid Composition of *Scoliodon laticaudus*

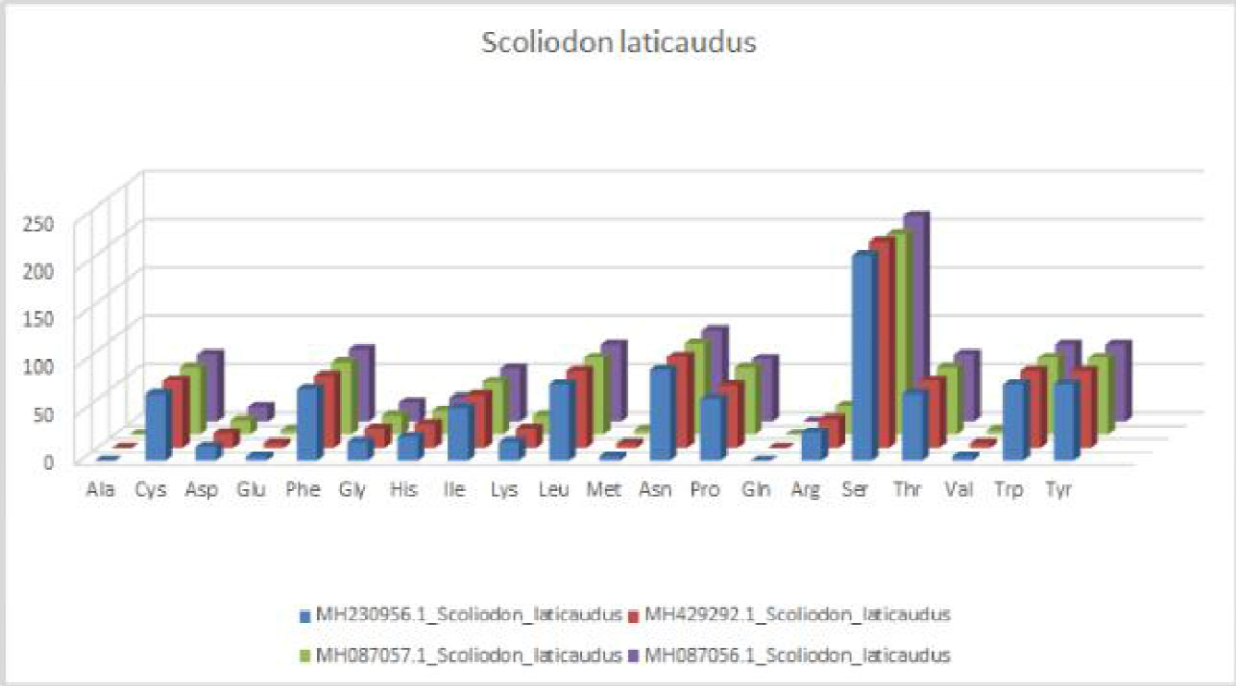

Figure 3.5.4: 20 amino acids percentage (Multiplied by 10) of *Scoliodon laticaudus*

Amino Acid Composition of *Rhizoprionodon oligolinx*

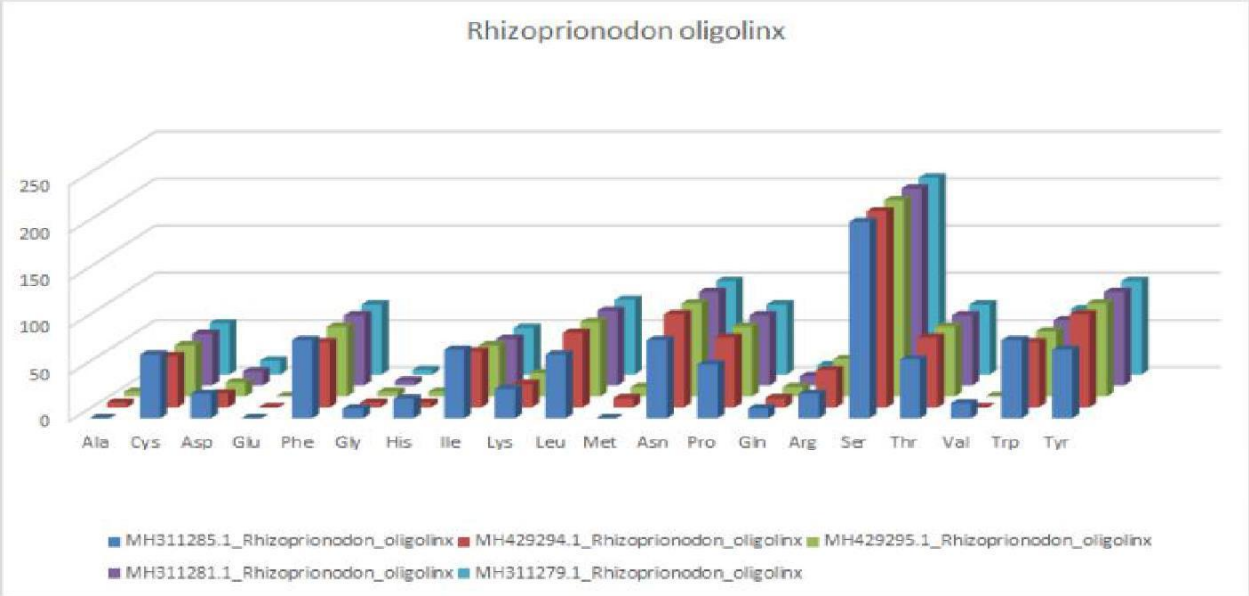

Figure 3.5.5: 20 amino acids percentage (Multiplied by 10) of *Rhizoprionodon oligolinx*

Amino Acid Composition of *Galeocerdo cuvier*

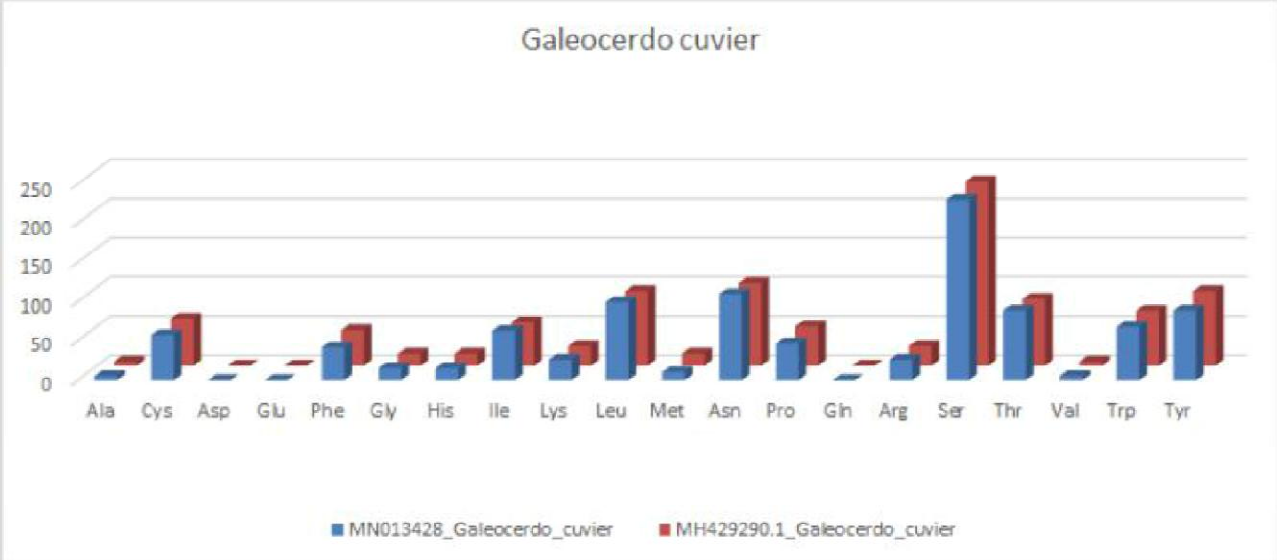

Figure 3.5.6: 20 amino acids percentage (Multiplied by 10) of *Galeocerdo cuvier*

Amino Acid Composition of *Narcine brunnea*

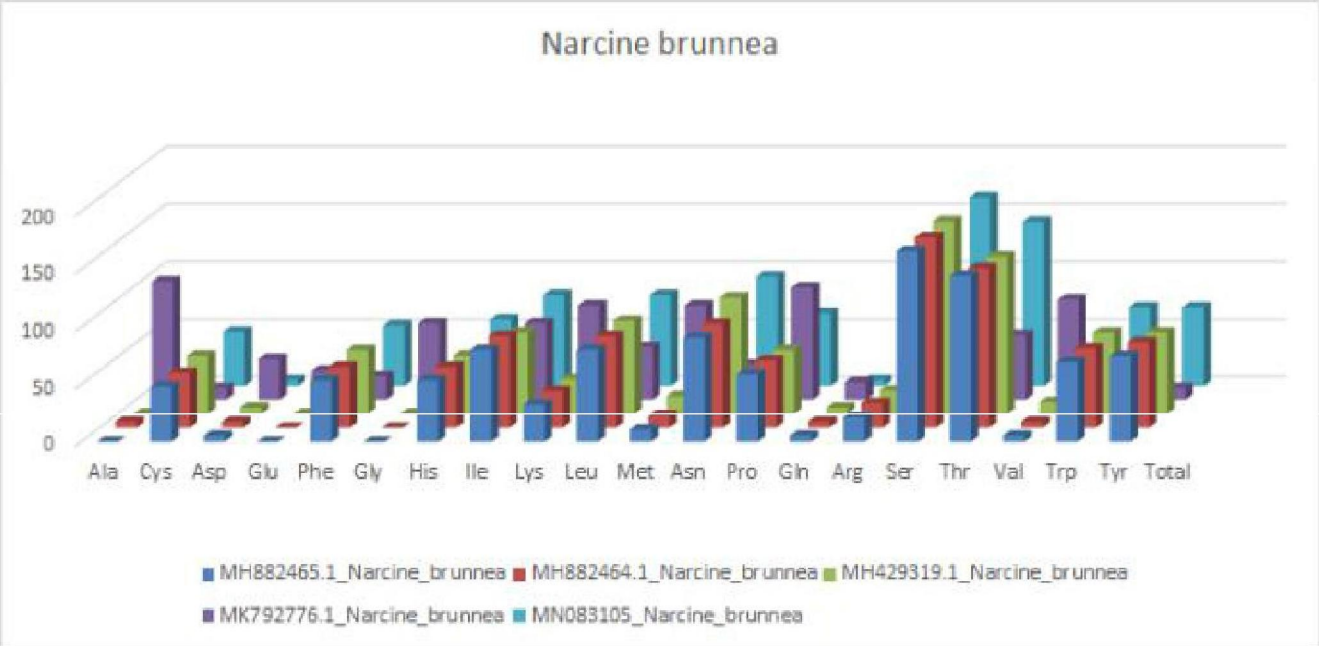

Figure 3.5.7: 20 amino acids percentage (Multiplied by 10) of *Narcine brunnea*

Amino Acid Composition of *Himantura Walga*

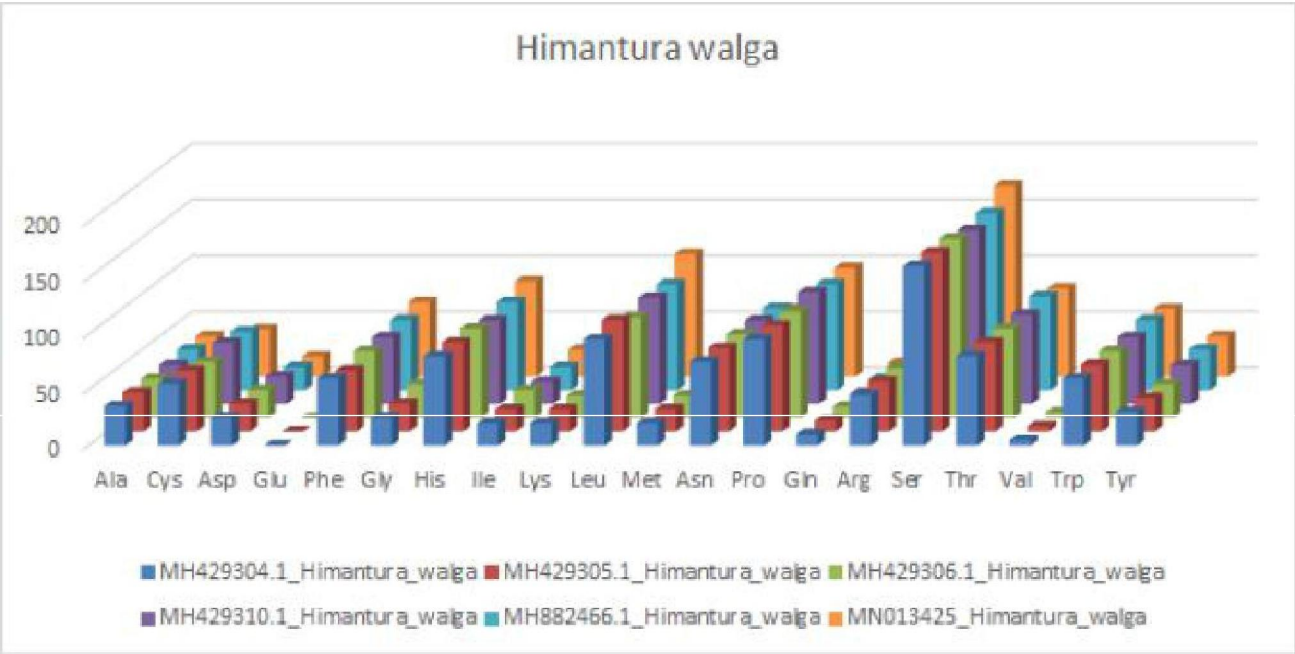

Figure 3.5.8: 20 amino acids percentage (Multiplied by 10) of *Himantura Walga*

Amino Acid Composition of *Neotrygon kuhlii*

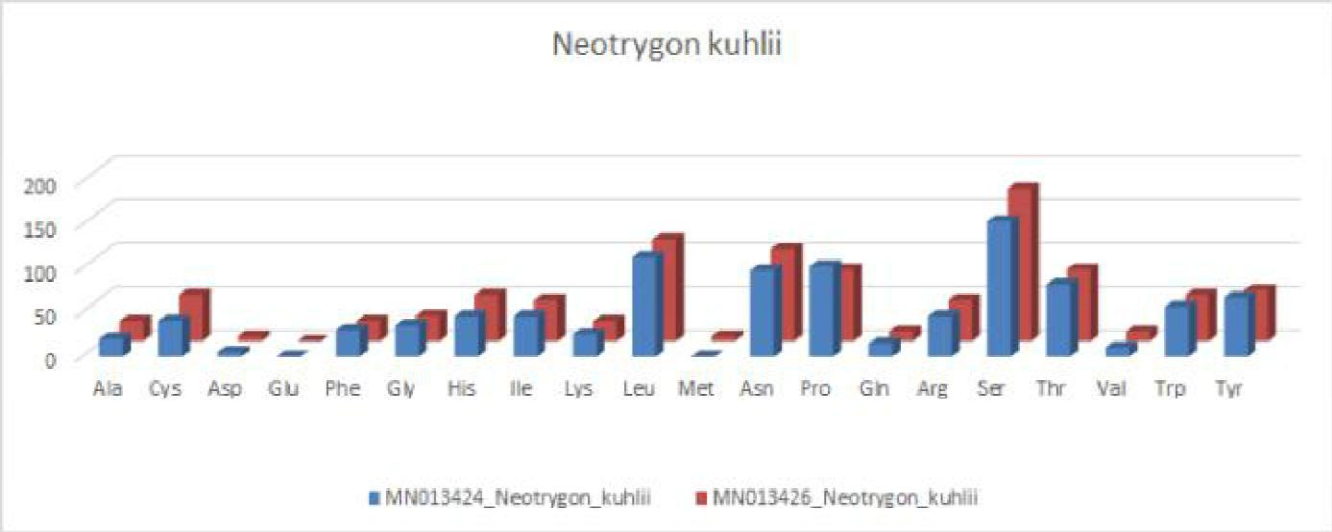

Figure 3.5.8: 20 amino acids percentage (Multiplied by 10) of *Neotrygon kuhlii*

Amino Acid Composition of *Gymnura poecilura*

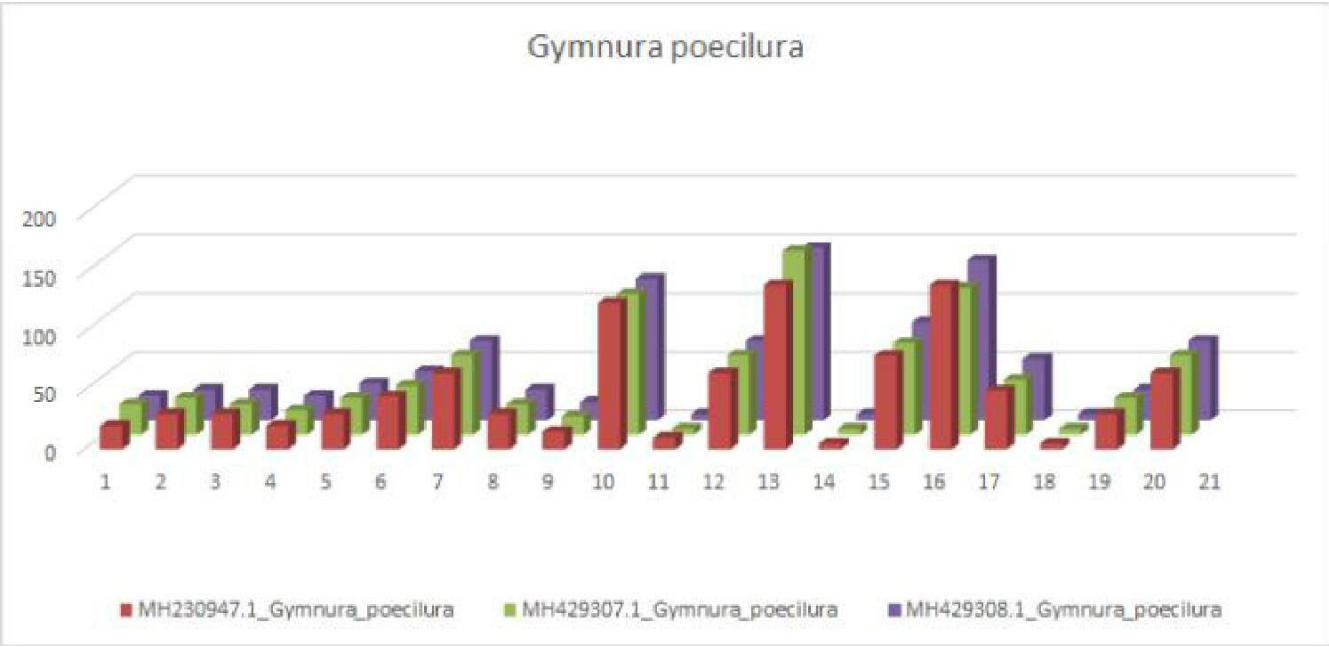

Figure 3.5.9: 20 amino acids percentage (Multiplied by 10) of *Gymnura poecilura*

Amino Acid Composition of *Pateobatis uarnacoides*

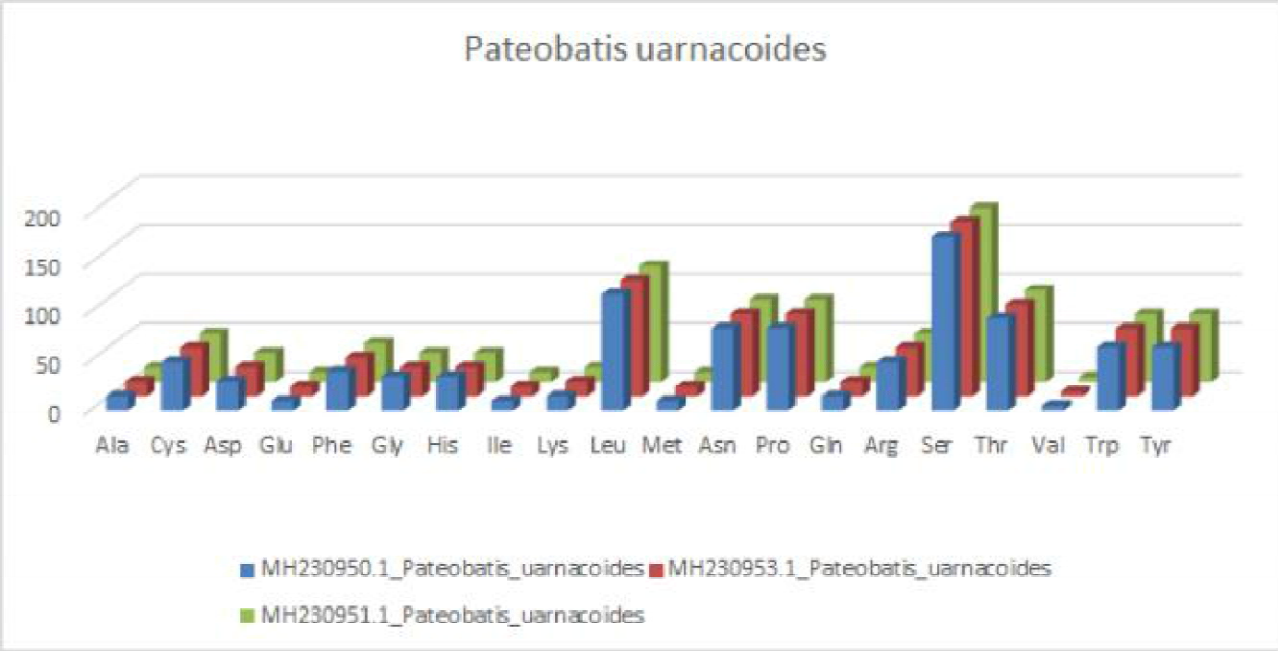

Figure 3.5.10: 20 amino acids percentage (Multiplied by 10) *Pateobatis uarnacoides*

3.5.1 Remarks

Each species wise Amino profiling shows similar pattern bars for different samples for same species which is a clear indication that amino acid profiling can be used for identification of species.

3.6 Cluster Bar representation of amino acid among shark species

With Cluster bar visual representation of 4 amino acid at a time we can see a clear pattern of change among different species of shark. This allows us to have a clear indication of the difference of each species and will help us to cross check two or more species even though their GC number comes out similar since the expression of protein might be different even though they have the same number of GC% [27]. The chart values of amino acid percentage are multiplied with ten for cluster bar visual aid.

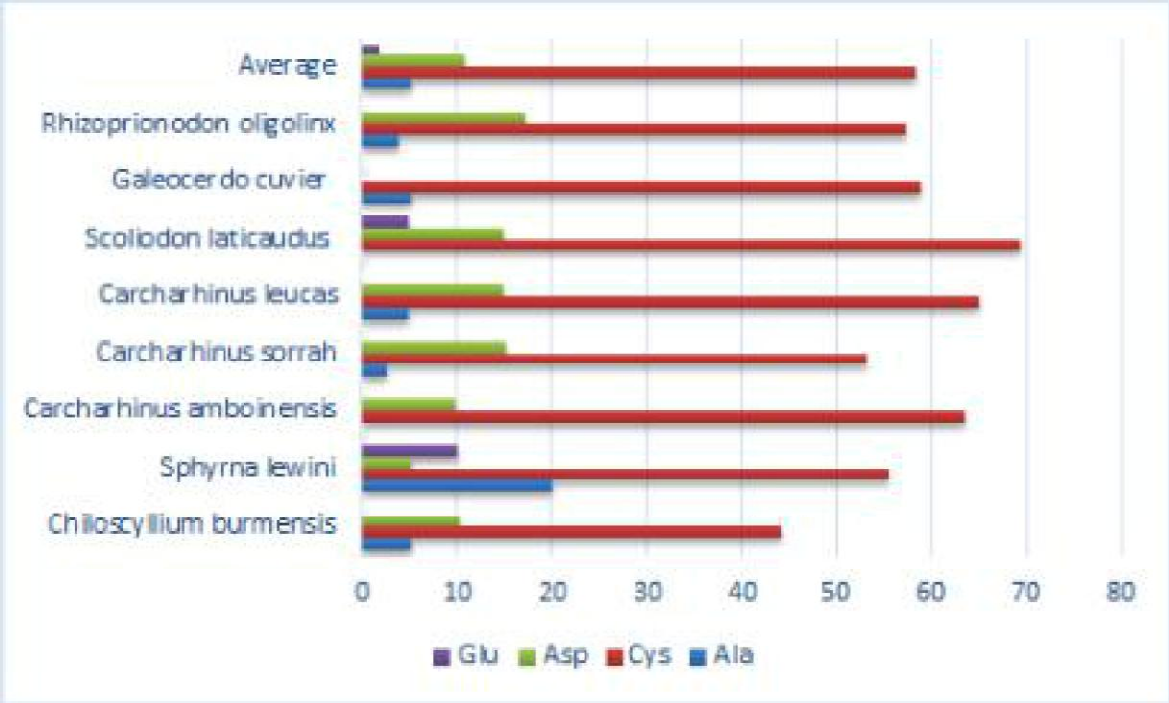

Figure 3.6.1: Cluster chart of 4 amino acid among 8 species of shark.

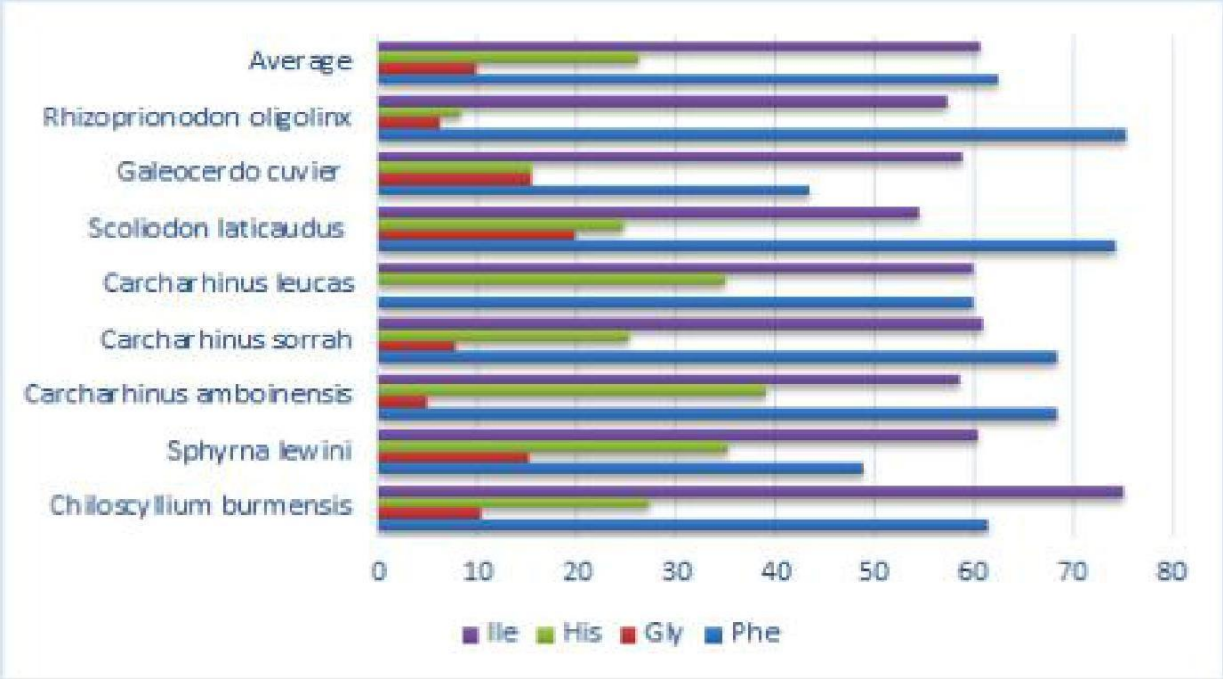

Figure 3.6.2: Cluster chart of 4 amino acid among 8 species of shark.

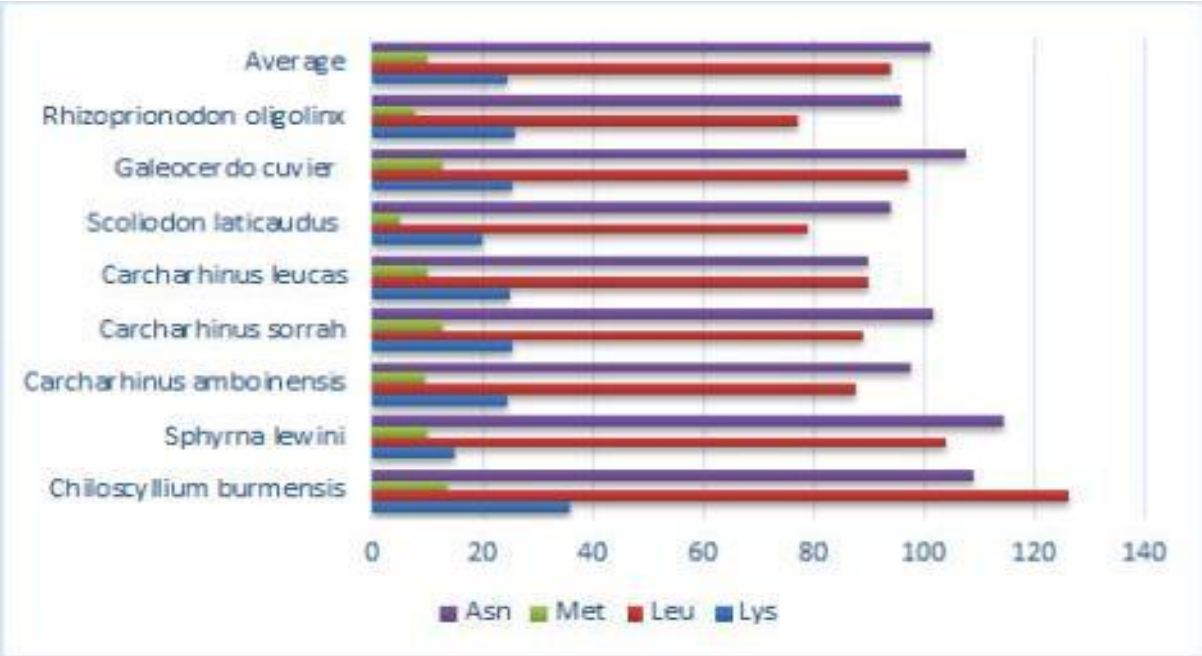

Figure 3.6.3: Cluster chart of 4 amino acid among 8 species of shark.

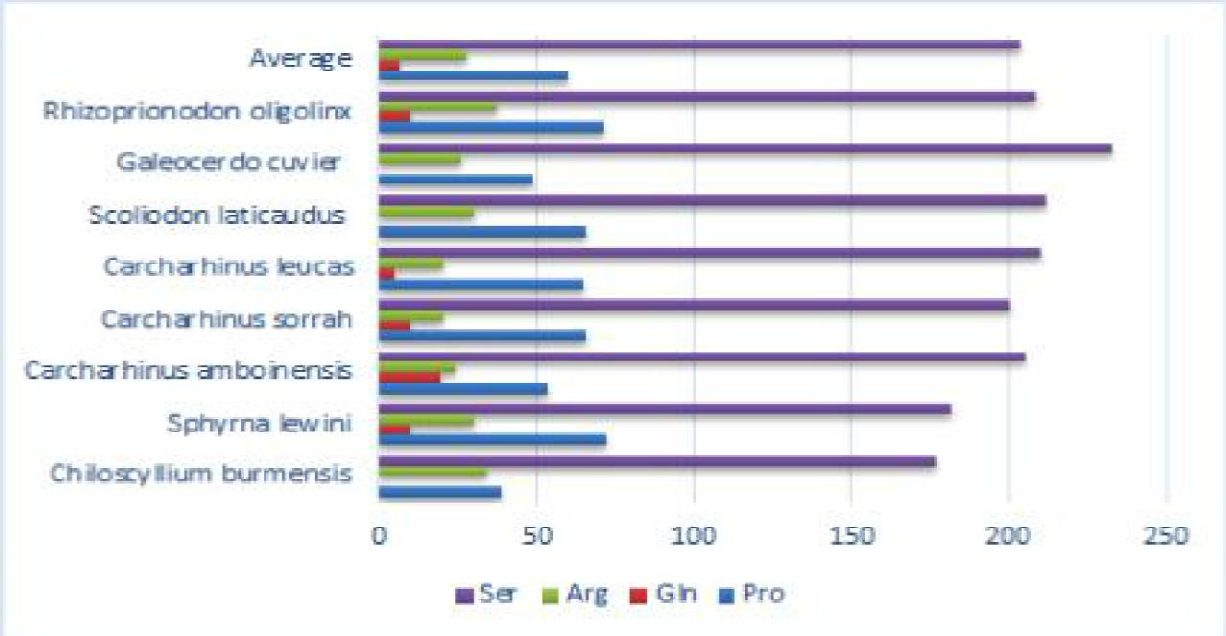

Figure 3.6.4: Cluster chart of 4 amino acid among 8 species of shark.

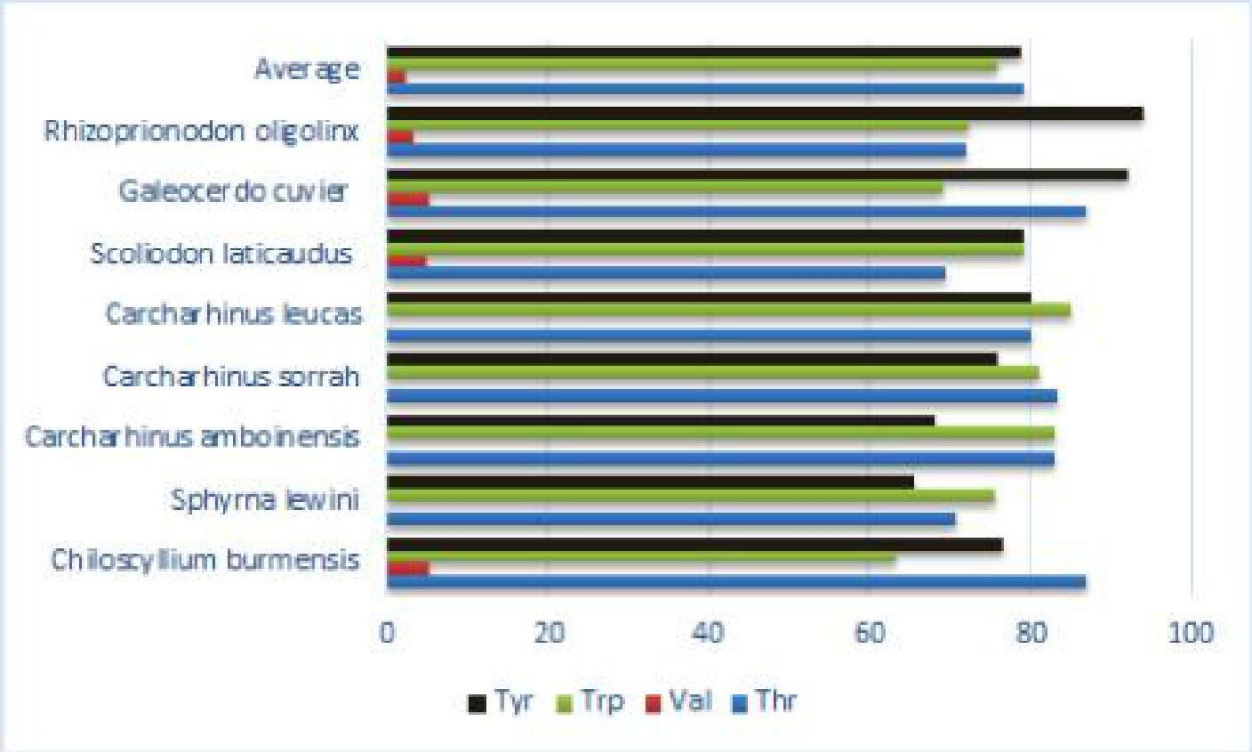

Figure 3.6.5: Cluster chart of 4 amino acid among 8 species of shark.

3.6.1 Remarks

When Amino Acid profiling is done between various species of sharks the range of amino acid in percentage varies even though the GC content may have been on a very close range and composition of nucleotide showed various similarities. Which means cluster bar representation can actually help to differentiate even closely related species. Significant values when amino acid percentages are viewed by species is that the lowest percentage of amino acid on average present in each species is Glutamic Acid at 0.187 % and the highest amino acid on average present in each species is Serine at 18.03 % (In chart each value is multiplied with 10 to make it more visible in cluster chart)

3.7.1 Variation of amino acid after multiple sequence alignment between shark species

Almost all shark species aligned in similar amino acid sequence pattern except *Rhizoprionodon oligolinx,* which had a point mutation at 120 number amino acid showing Methionine (M) which is Leucine (L) in the rest of the other species taken. [Table-10]

Table-10 Variation points of shark species

Species	Point of Amino Acid Variation	Common Amino acid/ Variation	Accession Number
Rhizoprionodon oligolinx	120	L/M	14.MH311285.1 15.MH429295.1 16.MH311281.1 17.MH311279.1

3.7.2 Variation of amino acid after multiple sequence alignment between ray species

When multiple sequence alignment was done between various species of rays, most of them were aligned in perfect sync but there was variation at various point of different species. The variation among different species may be different but the samples taken in consideration has the same sort of variation point species wise, which gives more solid ground to claim of showing variation point among species of shark and ray. [Table-11]

Table-11 Variation points of ray species

Species	Point of Amino Acid Variation	Common Amino acid/ Variation	Accession Number
Narcine brunnea	29 156 158 159 160 176	A/T P/S A/S I/V S/T V/I	1. MH882465.1 2. MN083105 3. MH882464.1 4. MH429319.1
Narcine malculate	29 119 156 158 160 176	A/T L/I P/S A/S S/T V/I	1. MN083137
Pateobatis uarnacoides	129	T/A	1. MH230953.1 2. MH230951.1 3. MH230950.1
Himantura undulata	129	T/A	4. MN013427
Himantura walga	129 159	T/A I/L	1. MH429310.1 2. MH429305.1 3. MH429304.1
Himantura gerrardi	129	T/A	1. MH230945.1
Mobular mobula	175	A/T	1. MH230952.1
Gymnura poecilura	175 176	A/T V/I	1. MH230947.1 2. MH429307.1 3. MH429308.1
Neotrygon kuhlii	175	A/T	1. MN013424 2. MN013426
Pateobatis jenkinsii	186 188	A/T G/D	1. MH230946.1

3.8 Phylogenetic Tree of Sharks and Rays

3.8.1 Phylogenetic analysis of sharks (2 Orders)

8 species of sharks were taken to create phylogenetic tree to find an ancestor connection between species of sharks, find closely related species, how far the drift across the species and order. [43]

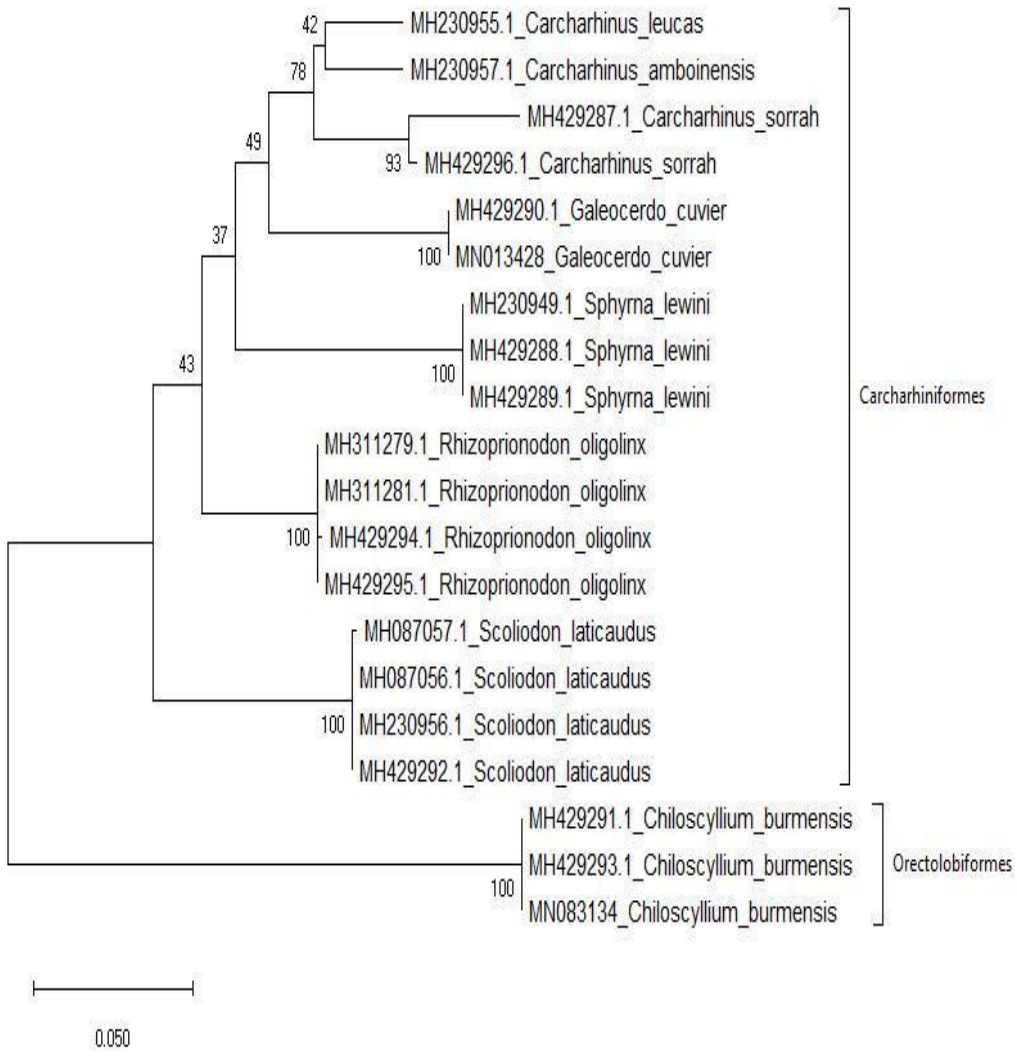

Figure 3.8.1.1: Original tree Constructed with 1000 replication and the method of construction is Maximum Likelihood of two shark Order with MEGA-X: Molecular Evolutionary Genetics Analysis in Windows operation system.

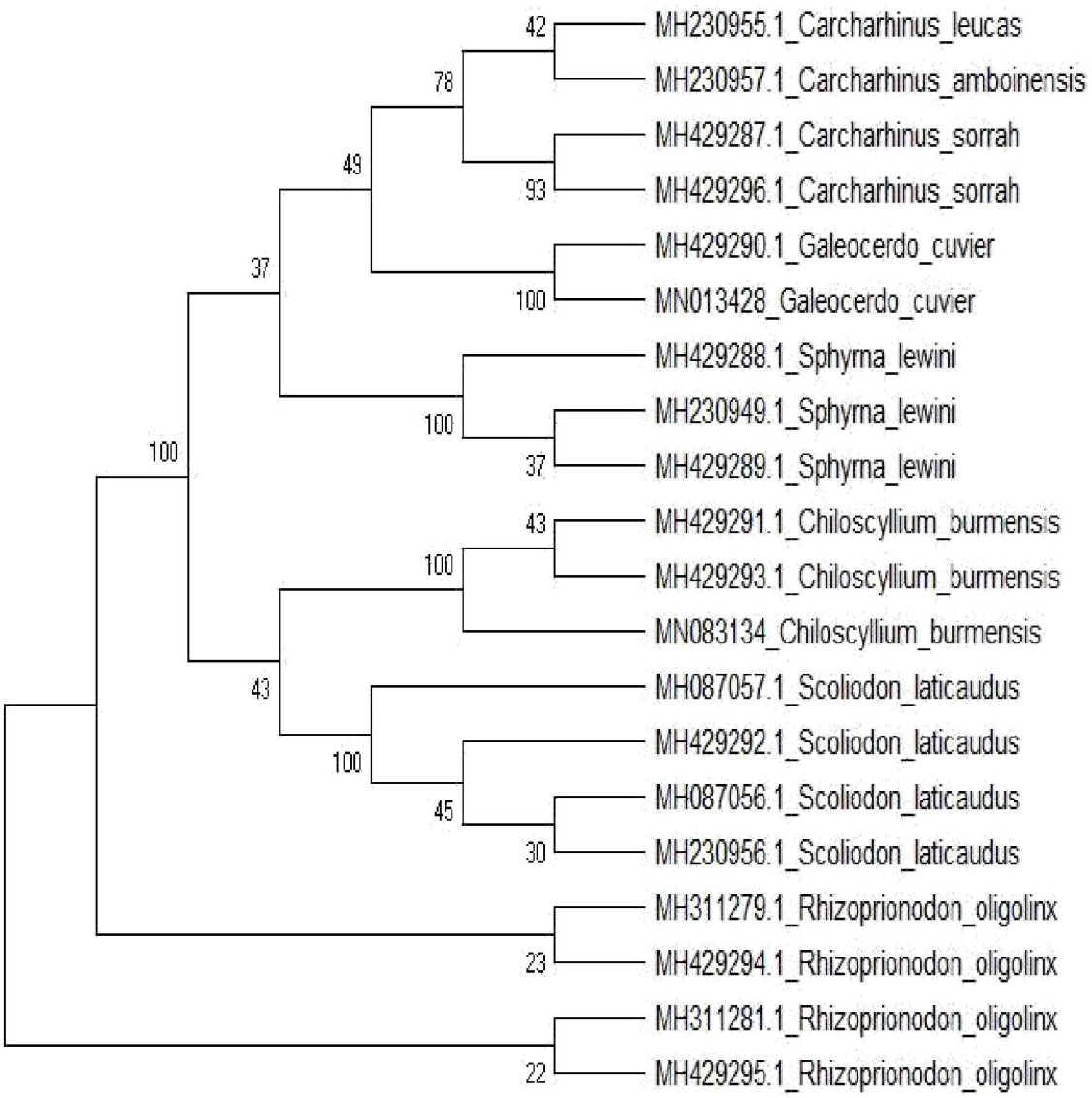

Figure 3.8.1.2: Bootstrap Constructed with 1000 replication and the method of Contruction is Maximum Likelihood of various shark species with MEGA-X: Molecular Evolutionary Genetics Analysis in Windows operation system.

3.8.2 Phylogenetic analysis of rays (2 Orders)

10 species of rays were taken to create phylogenetic tree to find an ancestral connection between species of rays, find closely related species, how far the drift across the species and order. [43]

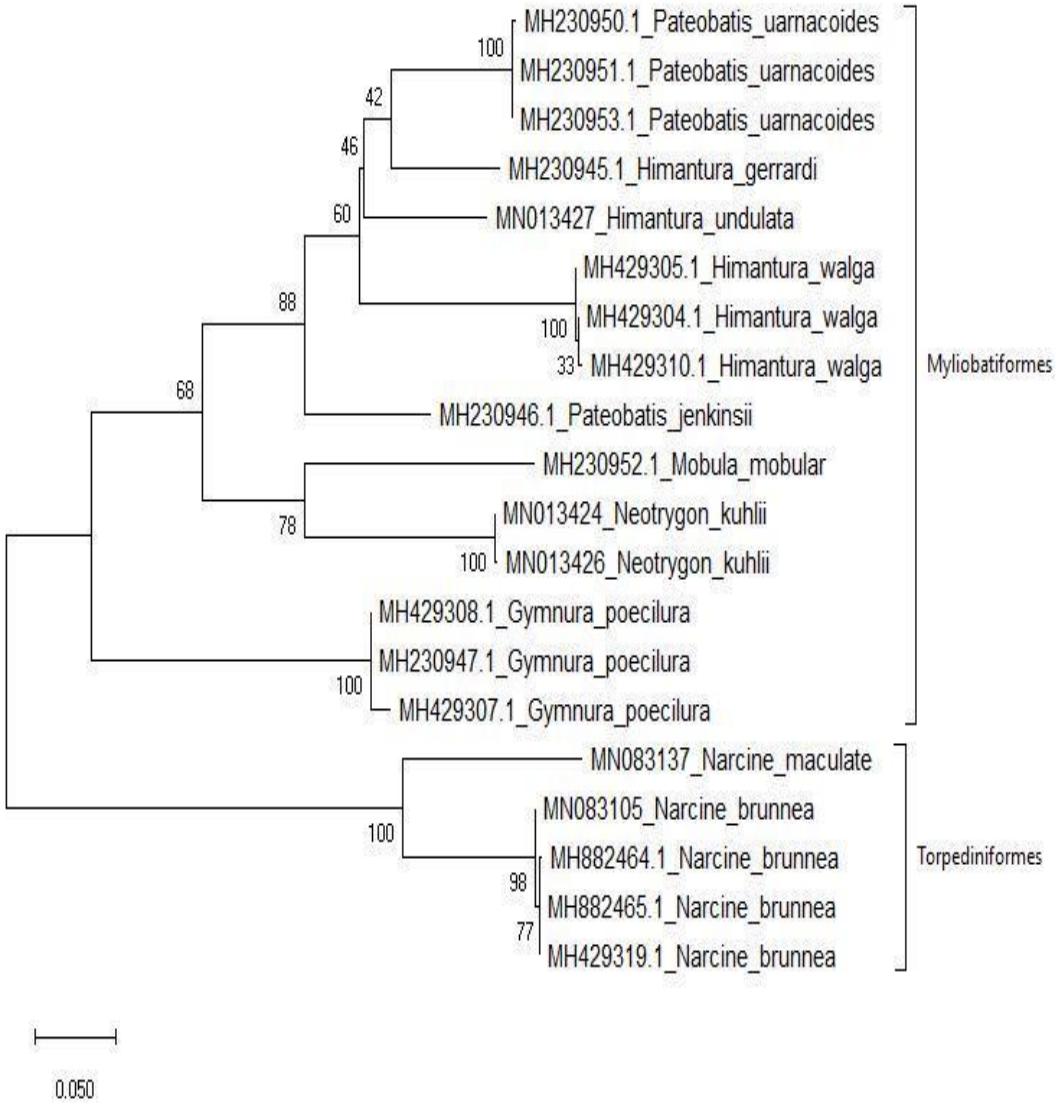

Figure 3.8.2.1: Original tree Constructed with 1000 replication and the method of construction is Maximum Likelihood of rays with MEGA-X: Molecular Evolutionary Genetics Analysis in Windows operation system.

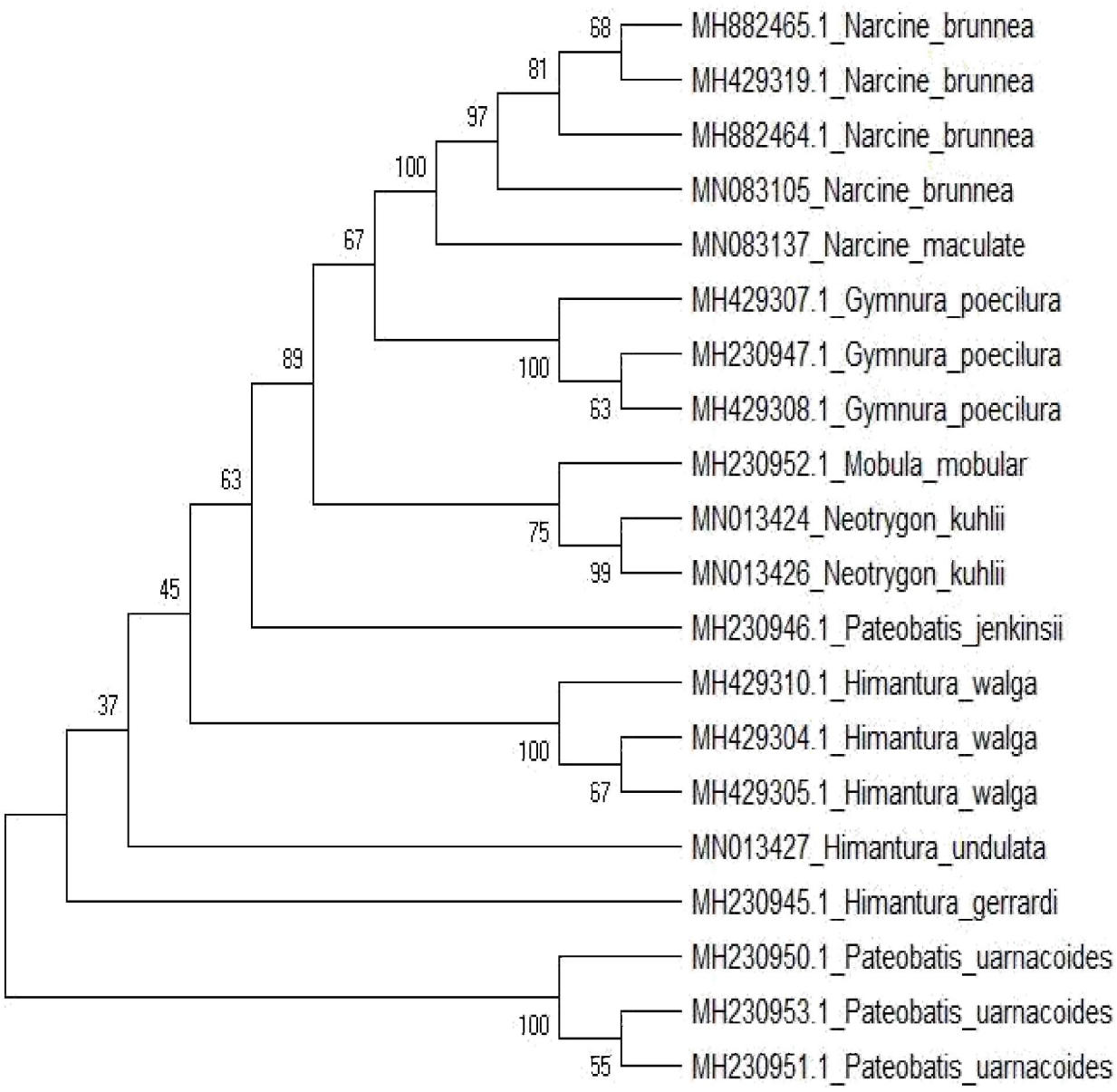

Figure 3.8.2.2: Bootstrap Constructed with 1000 replication and the method of Contruction is Maximum Likelihood of various ray species with MEGA-X: Molecular Evolutionary Genetics Analysis in Windows operation system.

3.8.3 Phylogenetic analysis of sharks and rays

A total number of 18 species of sharks and rays with 4 orders were taken to create phylogenetic tree showing the ancestral connection between them. [43]

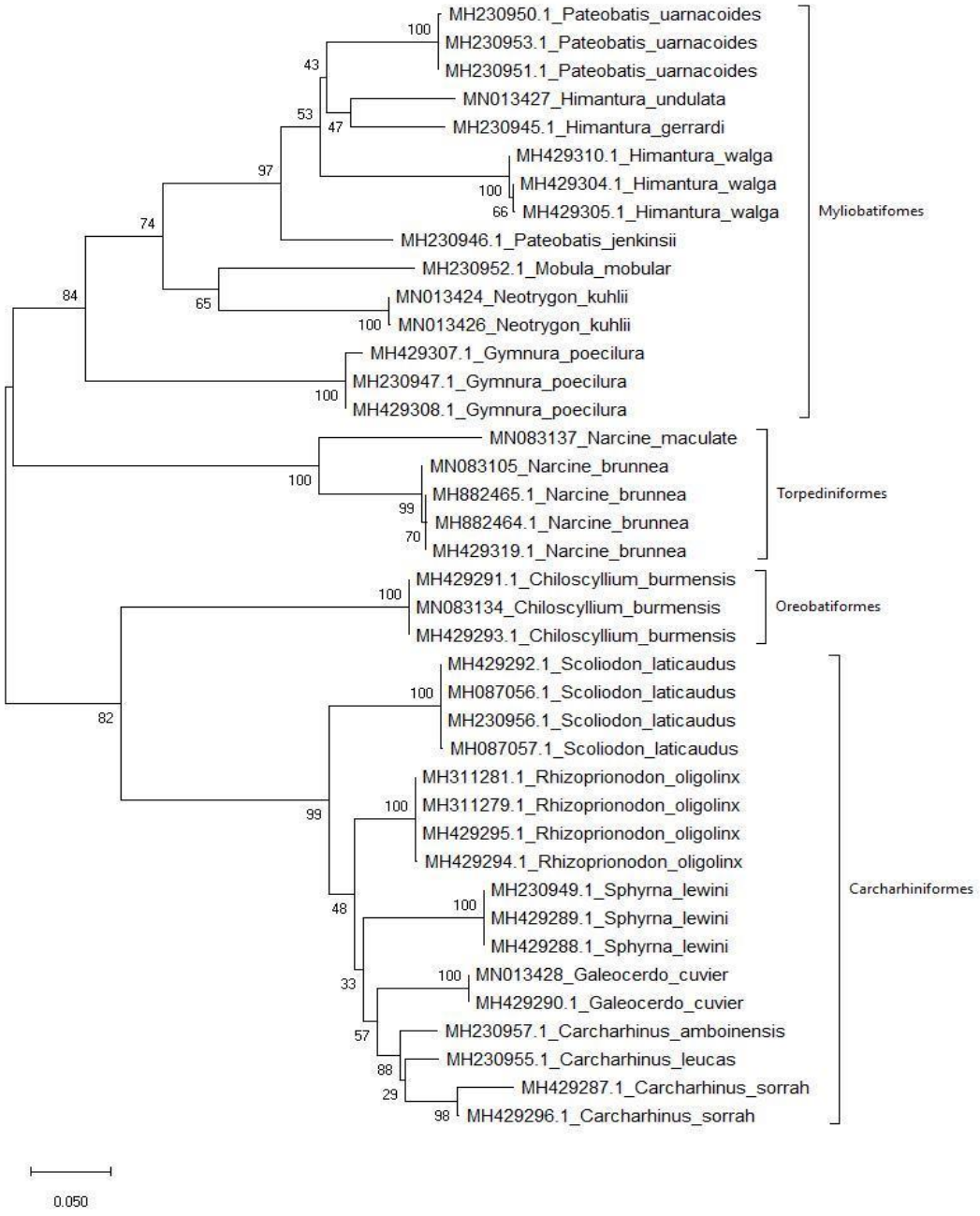

Figure 3.8.3.1: Original tree Constructed with 1000 replication and the method of construction is Maximum Likelihood of sharks and rays with MEGA-X.

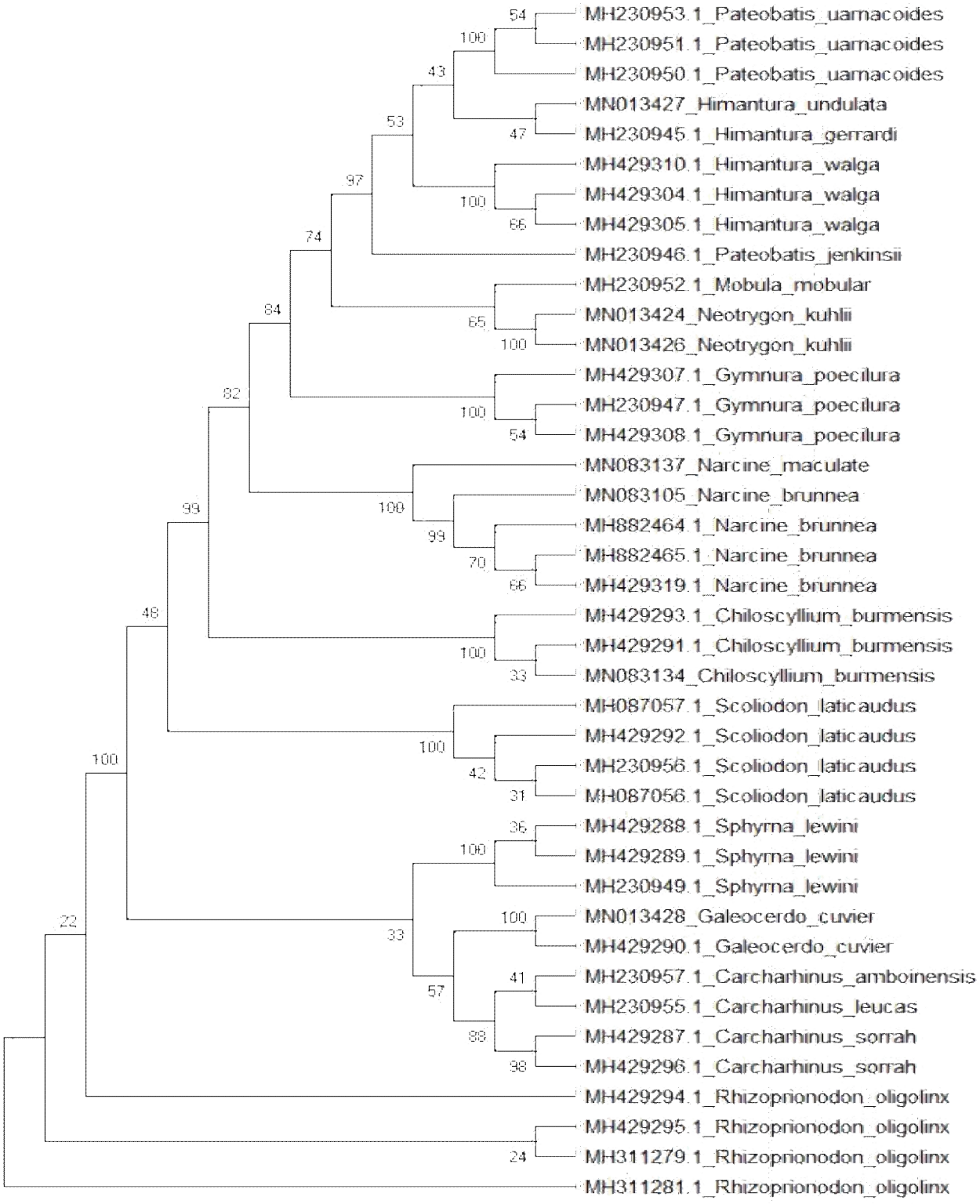

Figure 3.8.3.2: Bootstrap Constructed with 1000 replication and the method of Contruction is Maximum Likelihood of various sharks and ray species with MEGA-X.

Chapter 4 Discussion

DNA Barcoding has become a common tool for any sort of animal which came as a collective initiative from the biologist community to help them speak the same language in terms of identification and the core of the model is based on a universal language of DNA nucleotide sequence [28]. In this study we tried to enrich that database with various sharks and rays species sequences and analyzing them to check in the data matches the usual patterns and if the sequences are properly labeled to be used as future reference.

The choice of Mitochondrial COI gene is based on the fact that this region of the gene is usually the most conserved region of any species of animal and hence it is ideal for the purpose of identification. We on average took 655bp of nucleotide sequence for each sample of species. Even though this region of the gene is supposed to be conserved the variation can be seen in many sequences of same species which can be due to poor quality of sequencing process or frame shifting and mismatching of nucleotide in sanger sequencing process [29].

Even though GC content measurement is set as standard for DNA barcoding identification of species, in this study we found a strong connection with amino acid profiling and further analysis can be done to see where the mutation of protein is happening with protein mapping by determining where to look for variation with our primary assessment of the amino acid percentage variance of within species and between species [30].

Among the 8 species of shark taken in consideration for this study, *Chiloscyllium burmensis* is listed in IUCN as Data Deficient species. *Carcharhinus amboinensis, Carcharhinus sorrah, Carcharhinus leucas, Scoliodon laticaudus,* and *Galeocerdo cuvier* are identified as Near Threatened species. *Rhizoprionodon oligolinx* is listed as Least Concerned species but the most interesting and most valuable find among the shark sample would be *Sphyrna lewini* which is recently been listed as endangered [31].

Among the 10 species taken in consideration for this study, *Narcine maculata, Neotrygon kuhlii* are listed as Data Deficient, *Brevitrygon walga,* and *Gymnura poecilura* as Near Threatened in IUCN redlist of threatened species. *Himantura undulata, Himantura gerrardi, Pateobatis uarnacoides,* and *Pateobatis jenkinsii* is listed as Vulnarable species. We have found one species

of rays to be listed as Endangered which is *Mobula mobular. Narcine brunnea* is not evaluated in the redlist of threatened species of IUCN [31].

GC content on average for shark was 38.59 ± 0.12 and for rays was 43.67 ± 0.28 which indicates an almost 4-5% difference in their GC percentage which puts a clear difference line while differentiating GC percentage wise. The value trend of codons GC_1, GC_2, and GC_3 the usual $GC_1 > GC_2 > GC_3$ in both sharks and rays and the values were for shark 52.52 ± 0.23, 42.89 ± 0.04, and 20.38 ± 0.31 respectively and for rays 53.96 ± 0.36, 43.48 ± 0.19, and 33.59 ± 0.44 respectively which gives us an indication that the GC_1 stays well above 50% margin and the gradually decrease from $1^{st} > 2^{nd} > 3^{rd}$ codon which indicates that the mutation rate is highest in the third codon. The same pattern was observed in various papers related to barcoding of fish, insects and other animals. The combine GC values of shark and ray where overall GC content is 41.41±0.187. The value trend of codons GC_1, GC_2, and GC_3 the usual $GC_1 > GC_2 > GC_3$ in combined sharks and rays and the values were for sharks and rays combined are 53.32±0.17, 43.22±0.082, and 27.72±0.41 respectively [24,36,37].

In case of K2P distance study with primary sequences, we have found a clear scoring difference in Intra and Inter species K2p distance. The average intra species distance K2P% was 2.1% with SEM ±0.0008 and average Inter species distance was 29.5% with SEM ±0.0005. The minimum Intra distance value is 1.1% and Inter distance value is 11%. The maximum Intra distance value is 6.4% and Inter distance value is 38%. The significant finding is the gap between Inter and Intra value which is also called the DNA Barcoding gap which is 4.5 which is derived by subtracting the maximum intra distance and minimum inter distance which is close range to usual distance among intra and inter in other papers. [38,39, Table-09, Figure 3.4.1]

In case of K2P distance study with both primary sequence and reference sequence combined, we have found a clear scoring difference in Intra and Inter species K2p distance. The average intra species distance K2P% was 1.3% with SEM ±0.0007 and average Inter species distance was 30.4% with SEM ±0.0009. The minimum Intra distance value is 0% and Inter distance value is 11%. The maximum Intra distance value is 5% and Inter distance value is 63%. Also, the average intra genus, family and Order distance K2P% were 6.06%, 9.5% and 18.7% with SEM ±0.014, ±0.018 and ±0.04 respectively. The average Inter species, genus, family and order distance were 30.4%,

30.5%, 30.6% and 36.2% with SEM ±0.0009, ±0.008, ±0.003 and ±0.012 respectively. [38,39, Table-10]. This shows a clear similarities between two sets of results here.

Variation in Amino acid sequence after sequence alignment is visible in both sharks and rays. Sharks species alignment showed variation in 120th number amino acid where common amino acid is Leucine (L) and mutated amino acid is Methionine (M) [Table-10].

Rays species on the other hand had many variation points of different species. *Narcine maculata* and *Narcine brunnea* had mutations at 29th number amino acid where common amino acid is Alanine (A) and mutated amino acid is Threonine (T), 156th number amino acid where common amino acid is Proline (P) and mutated amino acid is Serine (S),158th number amino acid where common acid is Alanine (A) where mutated acid is Serine(S), 160th number amino acid where common amino acid is Serine (S) and mutated amino acid is Threonine (T), and 176th number amino acid where common acid is Valine (V) and where mutated acid is Isoleucine (I). *Narcine brunnea* has an extra mutation at the 159th number amino acid is Isoleucine (I) and mutated acid is Valine (V) (Table-11). *Pateobatis uarnacoides, Himantura undulata, Himantura walga, Himantura gerrardi* all has a mutation at 129th number amino acid where common amino acid is Threonine (T) and mutated amino acid is Alanine (A). *Himantura walga* has an extra mutation in the 159th number amino acid where common acid is Isoleucine (I) and mutated amino acid is Leucine (L). *Mobular mobula, Gymnura poecilura,* and *Neotrygon kuhlii* these three have variation of amino acid at the 175th number of amino acid where the common amino acid is Alanine
(A) and mutated amino acid is Threonine (T) (Table-11). Only *Gymnura poecilura* has an extra mutation at the 176th number of amino acid where the common amino acid is Valine (V) and the mutated amino acid is Isoleucine (I). *Pateobatis jenkinsii* has mutation on 186th and the 188th number amino acid where the common amino acid are as following Alanine (A) and Glycine (G) and the mutated amino acid are as following Threonine (T) and Aspartic acid (D) [35]. This sort of variation at specific point can be set as a marker for each species identification [Table-11].

While checking for the protein percentage of the *Narcine brunnea* sample with accession number MK792776.1 showed an irregular percentage of amino acid than other compared samples [Figure 3.5.7]. Which led to an investigation of whether the sample is misidentified,and looking back to

the morphological identification with picture taken of the species sample it gave an indication that there can be a mix up between the labeling between *Narcine brunnea* and *Narcine timlei.* Further, BLAST of the FASTA was done to check and the match the percentage identity which showed 93.73% match to *Narcine timlei* which gave more base to the confirmation that this sample of species was misidentified. It indicated that where COI nucleotide sequence based identification is problematic or confusing it can be confirmed through translated amino acid profile of COI gene.

References

1. Islam, M. K., Rahaman, M., & Ahmed, Z. (2018). Blue Economy of Bangladesh: Opportunities and Challenges for Sustainable Development. *Advances in Social Sciences Research Journal*, *5*(8).

2. Dawnay, N., Ogden, R., McEwing, R., Carvalho, G. R., & Thorpe, R. S. (2007). Validation of the barcoding gene COI for use in forensic genetic species identification. *Forensic science international*, *173*(1), 1-6.

3. Tillett, B. J., Field, I. C., Bradshaw, C. J., Johnson, G., Buckworth, R. C., Meekan, M. G., & Ovenden, J. R. (2012). Accuracy of species identification by fisheries observers in a north Australian shark fishery. *Fisheries Research*, *127*, 109-115.

4. Simpfendorfer, C. A., & Kyne, P. M. (2009). Limited potential to recover from overfishing raises concerns for deep-sea sharks, rays and chimaeras. *Environmental Conservation*, *36*(2), 97-103.

5. VALENTI, S., DULVY, N. K., FOWLER, S. L., CAVANAGH, R. D., HARRISON, L. R., CARLSON, J. K., & COMPAGNO, L. J. (2014). Extinction risk and conservation of the world's sharks and rays.

6. Hoq, M. E., Haroon, A. Y., & Hussain, M. G. (Eds.). (2011). *Shark fisheries in the Bay of Bengal, Bangladesh: Status and potentialities* (p. 76). Support to Sustainable Management of the BOBLME Project, Bangladesh Fisheries Research Institute.

7. Hoq, M. E. (2011). Sharks-a threatened biodiversity in the coastal water of Bangladesh. *Shark Fisheries in the Bay of Bengal, Bangladesh: Status and Potentialities*, 61.

8. Goto, T., Shiba, Y., Shibagaki, K., & Nakaya, K. (2013). Morphology and ventilatory function of gills in the carpet shark family Parascylliidae (Elasmobranchii, Orectolobiformes). *Zoological science*, *30*(6), 461-469.

9.Chowdhury, Gawsia. (2009). Encyclopedia of Flora and Fauna of Bangladesh vol. 24 Marine Fishes.

10. Cutmore, S. C., Bennett, M. B., & Cribb, T. H. (2010). Staphylorchis cymatodes (Gorgoderidae: Anaporrhutinae) from carcharhiniform, orectolobiform and myliobatiform elasmobranchs of Australasia: Low host specificity, wide distribution and morphological plasticity. *Parasitology international*, *59*(4), 579-586.

11. White, W. T., & Sommerville, E. (2010). Elasmobranchs of tropical marine ecosystems. In *Sharks and their relatives II* (pp. 175-256). CRC Press.

12. Ahmed, M. S., Chowdhury, N. Z., Datta, S. K., & Zhilik, A. A. (2019). New Geographical Record of the Burmese Bamboo Shark, Chiloscyllium burmensis (Orectolobiformes:

Hemiscylliidae), from Bangladesh Waters. *Thalassas: An International Journal of Marine Sciences*, 1-4.

13. Frazzetta, T. H., & Prange, C. D. (1987). Movements of cephalic components during feeding in some requiem sharks (Carcharhiniformes: Carcharhinidae). *Copeia*, 979-993.

14. Pank, M., Stanhope, M., Natanson, L., Kohler, N., & Shivji, M. (2001). Rapid and simultaneous identification of body parts from the morphologically similar sharks Carcharhinus obscurus and Carcharhinus plumbeus (Carcharhinidae) using multiplex PCR. *Marine Biotechnology*, *3*(3), 231-240.

15. Zeiske, E., Theisen, B., & Gruber, S. H. (1987). Functional morphology of the olfactory organ of two carcharhinid shark species. *Canadian Journal of Zoology*, *65*(10), 2406-2412.

16. Dulvy, N. K., Fowler, S. L., Musick, J. A., Cavanagh, R. D., Kyne, P. M., Harrison, L. R.,& Pollock, C. M. (2014). Extinction risk and conservation of the world's sharks and rays. *elife*, *3*, e00590.

17. Novianto, D., Nugroho, A. F., & Zedta, R. R. (2016). COMPOSITION AND ABUNDANCE OF PELAGIC SHARK CAUGHT BY DRIFT GILLNET IN CILACAP OCEANIC FISHING PORT, INDONESIA..

18. Snelson, F. F., Mulligan, T. J., & Williams, S. E. (1984). Food habits, occurrence, and population structure of the bull shark, Carcharhinus leucas, in Florida coastal lagoons. *Bulletin of Marine Science*, *34*(1), 71-80.

19. Jun, W., Shuyuan, Q., Yuling, H., Shengyun, Y., Xiaochun, L., & Mingru, C. (1996). Feeding Habits of spadenose shark, Scoliodon laticaudus from southern coast of Fujian. *Journal of Oceanography in Taiwan Strait*, *15*(4), 400-406.

20. Simpfendorfer, C. A., Goodreid, A. B., & McAuley, R. B. (2001). Size, sex and geographic variation in the diet of the tiger shark, Galeocerdo cuvier, from Western Australian waters. *Environmental Biology of Fishes*, *61*(1), 37-46.

21. Baum, J., Clarke, S., Domingo, A., Ducrocq, M., Lamónaca, A. F., Gaibor, N Gaibor, R. Graham & Martinez-Ortiz, J. (2007). Sphyrna lewini. IUCN red list of threatened species. *T39385A10190088*.

22. Schaefer, J. T., & Summers, A. P. (2005). Batoid wing skeletal structure: novel morphologies, mechanical implications, and phylogenetic patterns. *Journal of Morphology*, *264*(3), 298-313.

23. Tang, W. C. "Chinese medicinal materials from the sea." *Chinese Medicine* 1.4 (1987): 571-600.

24. Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1462), 1847-1857.

25. Frézal, L., & Leblois, R. (2008). Four years of DNA barcoding: current advances and prospects. *Infection, Genetics and Evolution*, *8*(5), 727-736.

26. Steinke, D., Vences, M., Salzburger, W., & Meyer, A. (2005). TaxI: a software tool for DNA barcoding using distance methods. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1462), 1975-1980.

27. Pentinsaari, M., Salmela, H., Mutanen, M., & Roslin, T. (2016). Molecular evolution of a widely-adopted taxonomic marker (COI) across the animal tree of life. *Scientific Reports*, *6*, 35275.

28. Hebert, P. D., Cywinska, A., Ball, S. L., & Dewaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1512), 313-321.

29. Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol*, *3*(5), 294-9.

30. Ortí, G., Petry, P., Porto, J. I., Jégu, M., & Meyer, A. (1996). Patterns of nucleotide change in mitochondrial ribosomal RNA genes and the phylogeny of piranhas. *Journal of Molecular Evolution*, *42*(2), 169-182.

31. Sembiring, A., Pertiwi, N. P. D., Mahardini, A., Wulandari, R., Kurniasih, E. M., Kuncoro, A. W.& Carpenter, K. E. (2015). DNA barcoding reveals targeted fisheries for endangered sharks in Indonesia. *Fisheries Research*, *164*, 130-134.

32. John, B. A., Asrul, M. M., Arshaad, W. M., Jalal, K. C. A., & Sheikh, H. I. (2018). DNA Barcoding of Rays from the South China Sea. In *DNA Barcoding and Molecular Phylogeny* (pp. 229-244). Springer, Cham.

33. Marchis, D., Benedetto, A., Amato, G., Brusa, B., Squadrone, S., & Abete, M. C. (2013). A quantitative real time polymerase chain reaction approach for estimating processed animal proteins in feed: preliminary data. *Italian Journal of Food Safety*, *2*(1), e4-e4.

34. Gil, L. A. (2007). PCR-based methods for fish and fishery products authentication. *Trends in Food Science & Technology*, *18*(11), 558-566.

35. Ward, R. D., & Holmes, B. H. (2007). An analysis of nucleotide and amino acid variability in the barcode region of cytochrome c oxidase I (cox1) in fishes. *Molecular Ecology Notes*, *7*(6), 899-907.

36. da Silva, J. M., Creer, S., Dos Santos, A., Costa, A. C., Cunha, M. R., Costa, F. O., & Carvalho, G. R. (2011). Systematic and evolutionary insights derived from mtDNA COI barcode diversity in the Decapoda (Crustacea: Malacostraca). *PLoS one*, *6*(5), e19449.

37. Jalali, S. K., Ojha, R., & Venkatesan, T. (2015). DNA barcoding for identification of agriculturally important insects. In *New horizons in insect science: Towards sustainable pest management* (pp. 13-23). Springer, New Delhi.

38. Čandek, K., & Kuntner, M. (2015). DNA barcoding gap: reliable species identification over morphological and geographical scales. *Molecular ecology resources*, *15*(2), 268-277.

39. Osmundson, T. W., Robert, V. A., Schoch, C. L., Baker, L. J., Smith, A., Robich, G, & Garbelotto, M. M. (2013). Filling gaps in biodiversity knowledge for macrofungi: contributions and assessment of an herbarium collection DNA barcode sequencing project. *PloS one*, *8*(4), e62419.

40. Kress, W. J., & Erickson, D. L. (2008). DNA barcodes: genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences*, *105*(8), 2761-2762.

41. Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., & Yoshikawa, S. (1995). Structures of metal sites of oxidized bovine heart cytochrome c oxidase at 2.8 A. *Science*, *269*(5227), 1069-1074.

42. Strüder-Kypke, M. C., & Lynn, D. H. (2010). Comparative analysis of the mitochondrial cytochrome c oxidase subunit I (COI) gene in ciliates (Alveolata, Ciliophora) and evaluation of its suitability as a biodiversity marker. *Systematics and Biodiversity*, *8*(1), 131-148.

43. Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*, *35*(6), 1547-1549.