INTERNATIONAL JOURNAL OF ADVANCE RESEARCH AND DEVELOPMENT

(Volume3, Issue4)

Molecular characterization of marine fishes using ITS2 region of ribosomal DNA

Neha Singh
Awadhesh Pratap Singh University,
Rewa, Madhya Pradesh

Pragya Singh
Suresh Gyan Vihar University,
Jiapur, Rajasthan

Arti Saxena
Model Science College,
Rewa, Madhya Pradesh

Dr. Samta Shukla
Awadhesh Pratap Singh University,
Rewa, Madhya Pradesh

Ugam Kumari Chauhan
Awadhesh Pratap Singh University,
Rewa, Madhya Pradesh

CONTENTS

Introduction

Plectorhincus nigrus (Brown sweet lip fish) and *Plectorhinchus flavomaculatus* (Gold spotted sweet lip) of Haemulidae family, and *Liza parsia* (Gold spot mullet) of Mugilidae family are marine fishes found in coastal region of Indian oceans. Fishes are chiefly marine, some brackish and rarely freshwater. These are ornamental fishes. *P.nigrus* and *P.flavomaculatus* according to their common name have round lip like structure, while *L.parsia* is gold spotted.

Because of their specific distribution in world, these are great interest to ecologist and evolutionary biologist. Their occurrence at specific regions of river gives impulse to the concept of territory specific evolution.

To choose suitable molecular markers or gene regions for resolving evolutionary uncertainty or investigating phylogenetic relationship among the organisms at a certain categorical level is a very difficult process. Different forces have led to evolution of various molecular markers or gene regions with varying degree of sequence variation. Thus, appropriate molecular markers or gene regions could be chosen with ever caution to deduce true phylogenetic relationship over a broad taxonomic spectrum. Nuclear organizer region has been used commonly by scholars and researchers for resolving taxonomic ambiguity **(Hwang and Kim, 1999).**

The presences of large variable regions in rDNA suggest the possibility of inter species variability in the sequence of these regions. The rDNA transcription tracts have low rate of polymorphism among species, which allows interspecific comparison to elucidate phylogenetic relationship using only a few specimens. Coding regions of rDNA is highly conserved among species but ITS regions are variable due to insertions, deletions, and point mutations. Low polymorphisms between rDNA repeat units, indicating that rDNA tandem arrays are evolving through concerted evolution.

Vertebrate ribosomal genes have several features that make them attractive to study phylogenetic issues. They include:

1) Highly conserved gene arrangement

2) Highly variable non - coding sequences.

3) Moderately repetitive gene copy number (**Lewin**, **1983**).

The analysis document that a huge amount of sequences variation exists in rDNA and thus the rRNA population is very heterogeneous. Although the true extent of inter-species and intra-individual variation is still largely unknown, it appears possible that no two rDNA transcribed sections are identical if one considered all the possible permutation and combinations. It remains to be determined how this diversity generated under pressure to homogenize the gene family.

Studies of phylogenetic become much valuable after use of molecular parameters. It is also very helpful to the DNA Barcoading.

Selection of molecular markers or gene resolving evolutionary uncertainty or investigating phylogenetic relationship at a certain categorical level is very difficult process. Different selective forces have led to the evolution of various molecular markers or gene regions with varying degree of sequence variation. Thus, appropriate molecular markers or gene region could be chosen with ever greater caution to deduce true phylogenetic relationship over a broad taxonomic spectrum. For the present study, I have analyzed ribosomal DNA sequences; form my experimental fish species with following objectives:

1. To study the genetic similarity among the species.
2. To establish phylogenetic relationship among the studied fishes.
3. To estimate the content of nucleotides in these species i.e. A, T, G, and C in percentage.

Knowledge of the correct phylogenetic relationship among animals is crucial for valid interpretation of evolutionary trends in biology.

Every living organism contains DNA, RNA, and Proteins. Closely related organism generally have a high degree of similarity in the molecular structure of these substances, while the molecule of organism distantly related usually show dissimilarity, but conserved sequences shows somewhat similarity even in distant species.

Conserved sequences such as ribosomal DNA are expected to accumulate mutations over time, and assuming a constant rate of mutation provides a molecular clock for dating divergence. Molecular phylogeny uses data to build a "relationship tree" that show probable evolution of various organisms. Not until recent decades, however it has been possible to isolate and identify these molecular structures.

The most common approach is the comparison of nucleotide sequence using sequence alignment technique to identify similarity. Another application of the phylogeny is in the DNA bar-coding where the species of an individual organism is identified by using small sections of mitochondrial DNA.

The effect on traditional biological classification scheme in the biological science has been dramatic as well. Work that was immensely labour and material intensive can now be done quickly and easily, leading to yet another source of information becoming available for systematic and taxonomic appraisal. This particular kind of data has become so popular that taxonomical schemes based solely on molecular data may be encountered. There are several biochemical data that can be used to infer phylogenetic relationship among species. Allozyme, immunological and DNA-DNA hybridization data have been widely used but are now increasingly being replaced by several types of DNA based data. Since that advent of PCR in 1985 - 86 **(Mullis *et al*., 1986; Saiki *et al*., 1985; Wrishink *et al*., 1987)** our knowledge about DNA of fish increased dramatically.

Genomic information of vertebrate increasing is rapidly because of several complete or almost complete vertebrate genome projects (e.g. Human, Mouse, Rat, Fugu, Zebra fish and Medaka). This information is invaluable to comparative biologists. Combined with developmental data gathered from model systems, such as Xenopus , Mouse Chicken, Medaka, And Zebra fish, and sophisticated post genomic analysis, comparative genomic studies are starting to discover evolutionary mechanisms underlying the diversification of body forms and the increasing complexity of gene function of vertebrates.**(Holland., 1999)** and **(Shimeld** and **Holland., 2000).**

The correct interpretation of comparative biological data requires an evolutionary framework (i.e. a well-supported phylogeny). Incorrect hypothesis of evolutionary relationships might result in misleading interpretation about, for example, taxon – specific genome characteristics and trends in gene family evolution, such as *hox* gene cluster evolution **(L.Z. Holland *et al*., 2003).**

Phylogenetic and comparatives biological data are most important to interpret early animal evolution **(A.H. Knoll and S.B. Carroll., 1999)** or vertebrate evolution. Unfortunately, a sound and widely accepted phylogenetic hypothesis for major fish lineages is not yet available.

My work is an effort to do the molecular characterization of three marine fish species that would be valuable for determination of evolutionary relationships of fishes.

Review of literature

The detection of genetic variation at the species, stock and population level is of great importance for sustainable aquaculture practices. Genetic variations at species level help us to identify the taxonomic units and to determine the species distinctiveness. Variations at the population level can provide an idea about different genetic classes, the genetic diversity among them and their evolutionary relationship with wild relatives. The genetic variability on stocks within population is extremely useful together for the information on individual identity, breeding patterns, degree of relatedness and disturbances of genetic variation among them **(Schierwater *et al*., 1994).**

The development of molecular techniques for genetic analysis has led to a great increase in our knowledge of genetics and our understandings of the structure of behaviour of genomes. These molecular techniques, in particular the use of the molecular marker, has been used to monitor DNA sequence variations in and among the species and create new sources of genetic variation by introducing new and favourable trait from landraces and related species.

Recent developments in molecular biology and bioinformatics have opened the possibilities of identifying and using genomic variation and major genes for the genetic improvement of organisms. The polymorphic DNA markers that were shown to genetically link to a trait of interest could be used for individual and pedigree identification, and trait improvement in genetics and breeding programmes. Morphological criteria **(Bernard and Talley., 2000),** biochemical data **(Jenen., 2000),** Isozyme electrophoresis **(Smith *et al*., 1997; Begg et al., 1998; Cagigas *et al*., 1999),** Restriction Fragment Length Polymorphism (RFLP) **(Hallerman and Bekmann., 1988; Garcia-Mas *et al*., 2000**), minisatellites **(Taggart and Ferguson., 1990**)**,** microsatellites **(Li *et al*., 2000**)**,** randomly amplified polymorphic DNA(RAPD) **(Welsh *et al*., 1991; Bernard and Talley., 2000; Fischer *et al*., 2000; Garcia-Mas *et al*., 2000; Lehmann *et al*., 2000**)

have been so far used to analyze genetic similarity and diversity in genetics and breeding research of animal.

Molecular phylogenetic relationships

Every living organism contains DNA, RNA and proteins. Closely related organisms generally have a high degree of agreement in the molecular structure of these substances, while the molecules of organisms distantly related usually show a pattern of dissimilarity. Conserved sequences, such mitochondrial DNA, are expected to accumulate mutations over time, and assuming a constant rate of mutation provides a molecular clock for dating divergence. Molecular phylogeny uses data to build a "Relationship Tree" that shows the probable evolution of various organisms. Not until recent decades, however, has it been possible to isolate and identify these molecular structures.

The most common approach is the comparison of nucleotide sequences using sequence alignment technique to identify similarity. Another application of molecular phylogeny is in DNA bar-coding; where the species of an individual organism is identify using small sections of mitochondrial DNA. Another application of the technique that makes this possible can be seen in the very limited field of human genetics, such as the ever more popular use of genetic testing to determine a child's, paternity as well as the emergence of a new branch of criminal forensic focused on evidence known as genetic fingerprinting. The effect on traditional biological classification schemes in the biological sciences has been dramatic as well. Work that was once immensely Labour and materials intensive can now be done quickly and easily, leading to yet another source of information becoming available for systematic and taxonomic appraisal. This particular kind of data has become so popular that taxonomical schemes based solely on molecular data may be encountered.

There are several kinds of biochemical data that can be used to infer phylogenetic relationships among species. Allozyme, immunological and DNA-DNA hybridization data have been widely used but are now increasingly being replaced by several types of DNA based data. Since the advent of

PCR in 1985-86 **(Mullis *et al*., 1986; Saiki *et al*., 1985; Wrishink *et al*., 1987)** our knowledge about DNA of fish increased dramatically.

PCR is an enzymatic cloning technique that allows the amplification of any stretch of DNA (with in size limits) that is flanked by synthetic oligonucleotide primers (**Saiki *et al*., 1985**). The primers usually around 20 bp in length and define 5' and 3' ends of the double stranded piece of DNA that is going to be amplified. The specificity of the amplification is accomplished through the need for an almost perfect fit of the primers to the template DNA. During each cycle of PCR the number of copies of the DNA fragment delineated by the primers at either end is doubled. Usually 25-40 cycles are completed in a thermal cycler in about 3 hours. PCR is much faster and economical than conventional cloning techniques.

First double stranded PCR product is produced that is then either sequenced (double stranded sequencing or alternatively cycle-sequenced) or sub cloned and then sequenced or cut with restriction enzymes or used as template DNA for subsequent asymmetric amplification **(Gyllensten and Erlich., 1988)** or digested with a exonuclease to produce single stranded DNA often allow one to read more base pairs than sequencing gels of double stranded DNA. Single strand PCR amplified DNA can be as clean as sub cloned DNA routinely more than 300-400bp can be unambiguously determined from a single sequencing reaction. The determination of nuclear DNA sequences will allow direct comparisons and study of DNA from different species.

The Phylogeny determined by ribosomal DNA

Molecular phylogenetic analysis begins with the sequence alignment. The role of sequence alignment is to organize sequences so that homologous residues appear in the same column of the alignment. It is important that sequence homology not be confused with sequence similarity. Homology implies common ancestry or of chance. All justification of phylogenetic inference rests on the assumption that the features compared are related by common ancestry. For molecular

phylogenies, this means that we hypothesize (or assert) that residues placed in a column of the alignment are descended from the same residue in a column ancestral gene. To the extent that our hypothesis is wrong, our conclusions lose justification. Thus, it is necessary to exclude sequence regions for which we cannot be confident of our alignment.

DNA sequence data play an essential role in the reconstruction of evolutionary relationships among the organisms, resulting in insights in genetic affinities that may confirm or conflict with traditional taxonomy. Because of attractive properties, the ribosomal DNA (rDNA) is popular for examine phylogenetic relationships and for studying genetic variability and divergence within and between species. Nuclear rRNA genes are found in multiple copies that are organized in tandem arrays, separated by transcribed spacers and surrounded by non-transcribed spacers along the genome. In eukaryotes, each repeat unit of the array consists of three rRNA genes (18S, 5.8S and 28S) and two internal transcribed spacers (ITS1 and ITS2) **(Long and David., 1980; Hillis and Dixon., 1991).** Each repeat of 28S rRNA gene is organized into several highly conserved cores interrupted by divergent regions, also called 'Domains' **(Hassouna *et al*., 1984).** The divergent domains evolve rapidly with substitution rates that are at least two orders of magnitude higher than those of core regions. These domains show a high rate of insertion and deletion events **(Olsen &Woese, 1993).** Variations in the more rapidly evolving ITS 2 region make them suitable for phylogenetic comparisons among closely related species.

Eukaryotes generally have many copies of the rDNA genes organized in tandem repeats; in humans approximately 300–400 rDNA repeats are present in five clusters (on chromosomes 13, 14, 15, 21 and 22

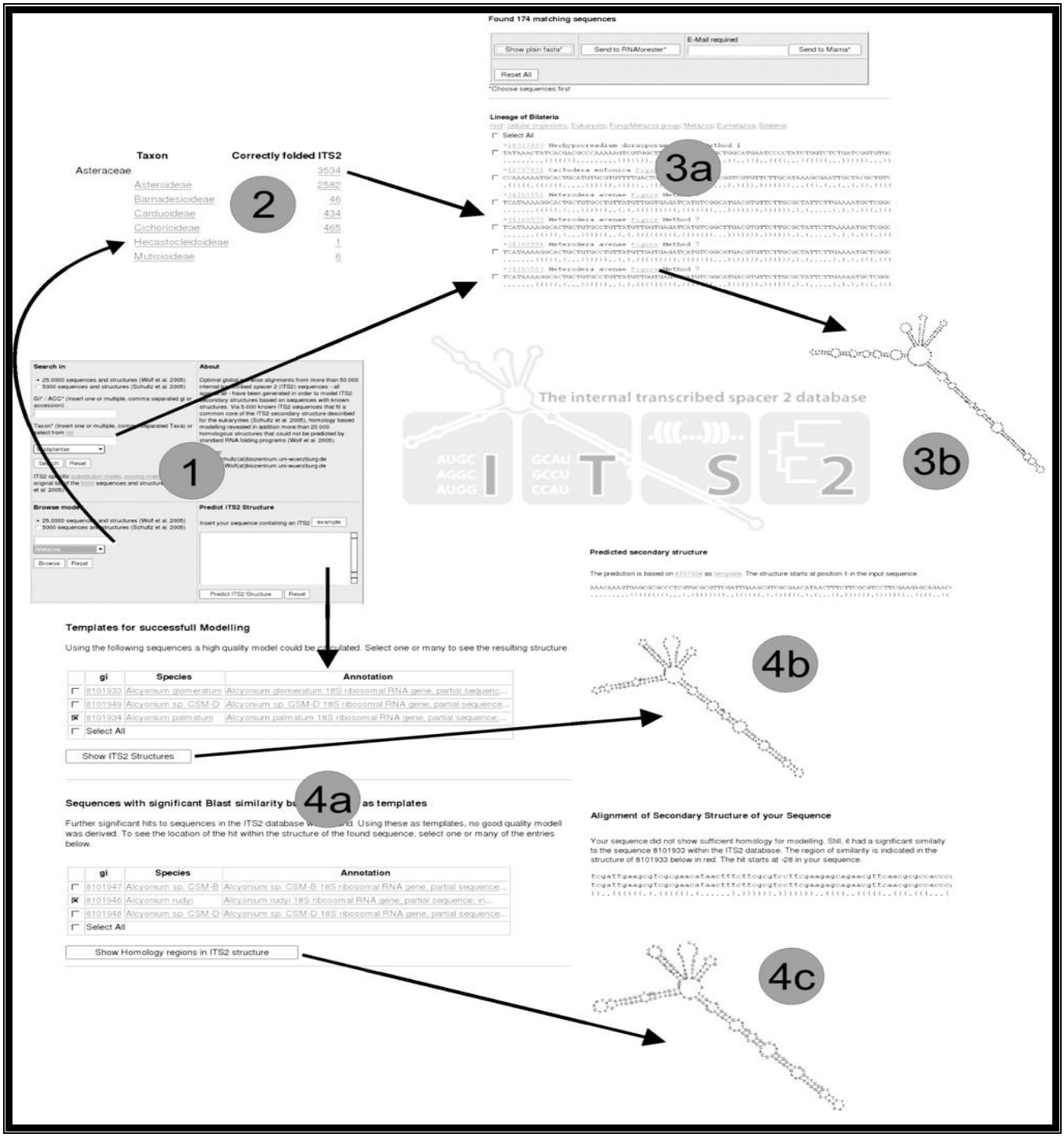


Fig 1-Screen shots of the web interface—(1) Start and Input view; (2) results of browsing for Asteraceae; (3a) list of ITS2 sequences and secondary structures in bracket notation for Bilateria, (3b) a sample secondary structure as stored in the database; (4) modelling of an ITS2 structure (4a) results of the BLAST search leading to high and low quality models, (4b) a predicted high quality model, (4c) homologous region indicated on the secondary structure of the template.

The prediction of the structure of a novel ITS2 will in most cases be the first step in a phylogenetic analysis. In a second step, the novel sequence has to be complemented with further sequences of the taxon of interest. Again, those with a known structure might be of highest use. We therefore implemented the ability to search for ITS2 sequences with correctly predicted structures (according to our model) not only within a species but within each taxon of interest. As a taxonomic classification, we are using the NCBI taxonomy database (with all its limitations according to phylogeny). In many cases, a taxonomic rank for the phylogenetic classification might not be known beforehand but depend on the coverage of the marker within the taxon. To provide a first overview, we implemented a browse mode, which allows traversing the NCBIs taxonomy (Figure 1). For each node and its siblings the number of ITS2 with predicted structures is given. Once the level of interest is reached, the ITS2 sequence and structure for all species belonging to this node can be retrieved. The resulting list of sequence-structure pairs can be the starting point for further analyses. To enable the upload to other, third party programs, the list can be exported in fasta format. A very common application will be the alignment of two or more sequences based on sequence and structure, in order to perform a phylogenetic analysis. Currently, there are only a few programs calculating global RNA alignments including sequence and structure information. Starting from the sequence-structure pairs of the ITS2 database, we allow to directly calling two of these programs, namely MARNA (10) and RNAforester (11) with selected sequences.

Already the ITS2 sequences combined with their secondary structures stored in the database might be of substantial interest for phylogenetic analyses. For example a recent analysis of the family Asteraceae (the daisy family) relied on 340 ITS2 sequences . To date, there are 3534 ITS2 sequences for this family with correctly folded structures within our database. A re-examination of this family considering these data might give a substantially broader coverage and thereby deeper insights into the evolution of this taxon. The collection of ITS2 together with the ability to predict the structure of novel sequences makes the ITS2 database a starting point for detailed phylogenetic analyses.

Therefore, it is of importance to make the interaction with further tools for phylogenetic analyses straightforward. On the one hand, we have already integrated access to MARNA and RNAforester , enabling the structure and sequence based alignment of ITS2 sequences. On the other hand, we provide a fasta formatted output which can be used as input for further programs. Once an alignment is calculated, standard phylogenetic programs like Phylip or more specific ones like the CBCanalyzer can be used for further analyses. In the near future, the ability to calculate sequence and structure based alignments as well as to edit them will be added to the web server (Seibel, P.N., Müller, T., Dandekar, T., Schultz, J. and Wolf, M., manuscript in preparation).

Molecular mechanisms of ITS2 processing, a eukaryotic insertion between the 5.8S and LSU rRNA, remain largely elusive even in yeast. To delineate ITS2 structural and functional features which could be common to eukaryotes, we first produced phylo-genetically supported folding models in the vertebrate lineage, then tested them in deeper branchings and, more particularly, among yeasts. ITS2 comparisons between four Teleostei, a Chondrichthyes specimen and two jawless organisms have revealed a common folding architecture in four to five domains of secondary structure emerging from a preserved structural core. This folding, largely reminiscent of ITS2 architecture in mammals, is also preserved in amphibia and in chicken, despite dramatic sequence variations. Preferential conservation is located around a central loop and at the apex of a long stem in the ITS2 3′-half. Interestingly, these two independent structural features contain, respectively, the 3′-ends of the two transient rRNA precursors 8S and 12S RNA identified in mammals, suggesting a preservation of these intermediates of processing over the entire vertebrate group. Surprising similarities between the vertebrate ITS2 folding shape and that of invertebrates as well as protista have made intriguing the significant differences from the yeast model. A detailed comparative analysis including four relatively close species and *Schizosaccharomyces pombe*, a deep yeast branching, has revealed an alternative phylogenetically supported four-domain folding presenting strong similarities to the vertebrate model. Remarkably, the two best conserved regions of vertebrates have unambiguously

preserved counterparts which are also sites for internal processing in yeast. Therefore, molecular mechanisms involved in ITS2 excision in vertebrates and yeast might be more closely related than currently believed and might require very similar *trans-acting* machinery.

ITS (for internal transcribed spacer) refers to a piece of non-functional RNA situated between structural ribosomal RNAs (rRNA) on a common precursor transcript. Read from 5' to 3', this polycistronic rRNA precursor transcript contains the 5' external transcribed sequence (5' ETS), 18S rRNA, ITS1, 5.8S rRNA, ITS2, 28S rRNA and finally the 3'ETS. During rRNA maturation, ETS and ITS pieces are excised and as **non-functional maturation by-products** rapidly degraded. Genes encoding ribosomal RNA and spacers occur in tandem repeats that are thousands of copies long, each separated by regions of non-transcribed DNA termed *intergenic spacer* (IGS) or *non-transcribed spacer* (NTS). Sequence comparison of the ITS region is widely used in taxonomy and molecular phylogeny because is (due to the high copy number of rRNA genes) easy to amplify even from small quantities of DNA, and b) has a high degree of variation even between closely related species. This can be explained by the relatively low evolutionary pressure acting on such non-functional sequences.

For example, ITS has proven especially useful for elucidating relationships among congeneric species and closely related genera in Asteraceae (Baldwin, 1992; Baldwin et al., 1995; Kim et al., 1996) as well as clinically important yeast species.

Importance of DNA based phylogenetic studies over Protein based analyses

The most sensitive comparisons between sequences are made at the protein level dedication of distantly related sequences is easier in protein translation, because the redundancy of the genetic code of 64 codons are reduced to 20 distinct Amino acids, the functional building blocks of proteins. However, the loss of degeneracy at this level is accompanied by a loss of information that relates more directly to the evolutionary process, because proteins are a functional abstraction of genetic events that occur in DNA. This is illustrated in the analysis of silent mutations, i.e. mutation at the DNA level that do not result in an amino acid substitution at the protein level (because of the redundancy of the genetic code), are automatically incorporated in to the analysis; these would not be detectable in a purely protein level treatment. **(Teresa K. Attwood., 2010)**

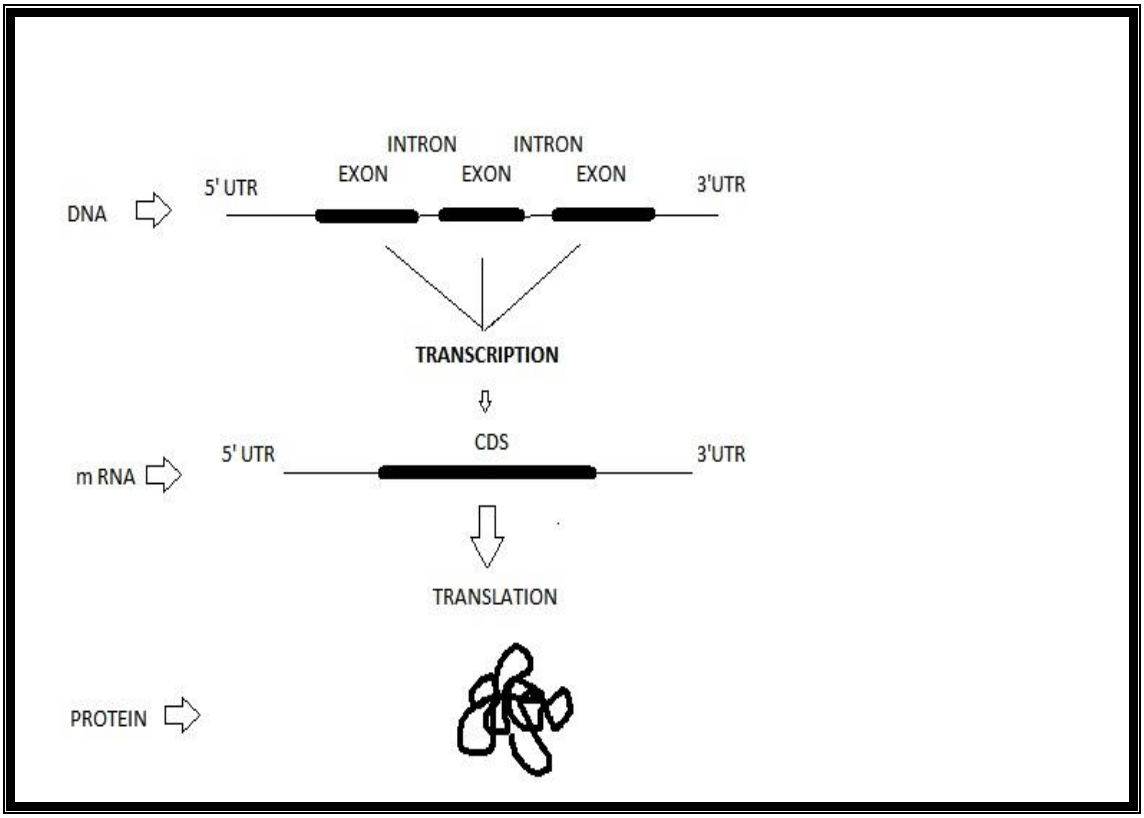


Fig-2 Showing the process of Transcription and Translation

Importance of rDNA

Ribosomal DNA characteristics are important in medicine and in evolution.

1- Transcription product of rDNA is the target of several clinically relevant antibiotics: Chloramphenicol, Erythromycin, Kasugamycin, Micrococcin, Paromomycin, Ricin, Sarcin, Spectinomycin, Streptomycin, and Thiostrepton.

2- The rRNA, usually present in all cells. For this reason, genes (rDNA) that encode the rRNA are sequenced to identify an organism's taxonomic group, calculate related groups, and estimate rates of species divergence. For this reason many thousands of rDNA sequences are known and stored in specialized databases such as RDP-II, SILVA and NCBI.

In summary, the described resource can be a starting point for any ITS2-based phylogenetic analysis and thereby complement databases for other phylogenetic markers like the European ribosomal database. Whereas the first might be of special use for low level analyses focussing on species and genus, the second might be more suited for higher level analyses. The combination of both will therefore give increased insight over a wide-range of taxonomic levels.

Materials and Method

Plectorhinchus niger

Kingdom : Animalia

Phylum : Chordata

Class : Actinopterygii

Order : Perciformes

Family : Haemulidae

Genus : *Plectorhinchus*

Species : *Nigrur*

Fig 3

The **sweet lips,** *plectorhinchus nigrus* in the family Haemulidae found in fresh, brackish and salt waters having Max length of about 45.0 cm. These fishes have big fleshy lips and tend to live on coral reefs in the Indo-Pacific in small groups or pairs. They will often associate with other fishes of similar species, and it is not unusual to see several species of sweet lips all swimming together. They are usually seen in clusters in nooks and crannies or under overhangs, The fish feed on invertebrates prey, such as bristle worms or shrimps and small crabs.

Plectorhinchus flavomaculatus

Kingdom : Animalia

Phylum : Chordata

Class : Actinopterygii

Order : Perciformes

Family : Haemulidae

Genus : Plectorhinchus

Species : *Flavomaculatus*

Fig 4

The *Plectorhinchus flavomaculatus (* Goldspotted Sweetlips) can be recognised by its colouration. It is found in tropical and some temperate marine waters of the Indo-West Pacific. Gold-spot Blubber-lips, Goldspot Sweetlip Bream, Lemon Sweetlips, Morwong, Netted Morwong, Netted Sweetlips, Rubberlip, Yellow-spotted Sweetlip .

The Goldspotted Sweetlips can be recognised by its colouration. It is grey to brown with numerous yellow to orange spots on the body, dorsal fin and anal fin. Juveniles have orange and grey-blue lines on the head and body that break into spots as a fish ages. The caudal fin usually has a dusky ventral margin. The species grows to 60 cm in length.

The species occurs in tropical and some temperate marine waters of the Indo-West Pacific. In Australia it is known from the south-western coast of Western Australia, around the tropical north and south to the southern coast of New South Wales.

Liza parsia

Kingdom : Animalia

Phylum : Chordata

Class : Actinopterygii

Order : Mugiliformes

Family : Mugilidae

Genus : Liza

Species : *Parsia*

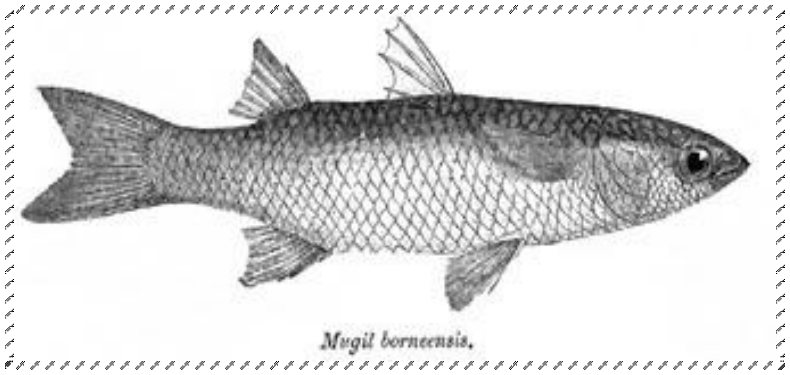


Fig 5

This particular species can be found coastal water,estuaries, lagoons , sometimes entering tidal rivers. The fish *Liza parsia* commonly known as **mullet** a family (muilidae) of ray- finned fish found worldwide in coastal temperate, tropical water and in some species in fresh water like the Mullets have served as an important source of food in various countries. Mullets are distinguished by the presence of two separate dorsal fins, small triangular mouths, and the absence of a lateral line organ. They feed on detritus small algae, diatoms and organic matter.

DNA ISOLATION

Deoxyribonucleic acid (DNA) is a nucleic acid that contains the genetic instructions used in the development and functioning of all known living organisms and some viruses. The main role of DNA molecules is the long-term storage of information. DNA is often compared to a set of blueprints or a recipe, since it contains the instructions needed to construct other components of cells, such as proteins and RNA molecules. The DNA segments that carry this genetic information are called genes, but other DNA sequences have structural purposes, or are involved in regulating the use of this genetic information.

Chemically, DNA consists of two long polymers of simple units called nucleotides, with backbones made of sugars and phosphate groups joined by ester bonds. These two strands run in opposite directions to each other and are therefore anti-parallel. Attached to each sugar is one of four types of molecules called bases. It is the sequence of these four bases along the backbone that encodes information. This information is read using the genetic code, which specifies the sequence of the amino acids within proteins. The code is read by copying stretches of DNA into the related nucleic acid RNA, in a process called transcription

DNA can be easily extracted from the muscle or blood of fish by Proteinase K, Phenol-chloroform-Isoamyl alcohol based protocol **(Russell and Sambrook 2001).** This is the most common method which removes proteins and other components from nucleic acids and relatively pure DNA can be obtained for further analysis.

REQUIREMENTS:

INSTRUMENTS

- Ultracentrifuge with angled rotor (Sigma)
- Water bath (set at 37^0 C) (Lab-Tech)
- Incubator (set at 50^0 C)
- Mettle balance
- Electrophoresis unit
- U.V. transilluminator
- U.V. Spectrophotometer
- Magnetic stirrer
- Mini centrifuge

GLASS WARES:

- Conical flasks
- Beakers
- Measuring cylinders
- Micro tubes (Eppendorf)
- Micropipettes
- Micro tips

CHEMICALS:

- Ethanol as preservative
- NaCl
- Tris-Cl
- EDTA
- 20% Sodium Dodecyl Sulphate
- Proteinase K
- Double distilled water
- Phenol
- Chloroform
- Isoamyl alcohol
- Ammonium acetate
- Isopropyl alcohol
- Milli Q water
- Bromo Phenol Blue
- Glycerol
- Sodium hydroxide
- Agarose
- Ethidium bromide

- TAE Buffer
- dNTPs
- 10X Taq Buffer
- MgCl2

REAGENTS:

a) 1M EDTA (pH 8.0)

Na_2 EDTA.$2H_20$ ---------- 31.22 g

Double distilled water ---------- 80.00 ml.

b) 1 M TRIS-HCl (pH 8.0)

Tris base ---------- 121.1 g.

Distilled Water ---------- 80.0 ml

pH = 8.0, final volume=100 ml and stored at 4^oC.

c) 1 M NaCl (pH 8.0)

NaCl (Salt) ---------- 5.85 g.

Distilled water ---------- 80.00 ml

pH =8.0, final volume =100 ml.

d) Sodium Dodecyl Sulphate (SDS) 20%

SDS	----------	20.00 g
Distilled water	----------	80.00 ml

Final volume=100 ml.

e) Proteinase K

10 mg. of Proteinase K dissolved in 1 ml of distilled water and store at -20^{o}C

f) Lysis Buffer

1 M EDTA (pH 8.0)	----------	376.00 µl
1 M NaCl (pH 8.0)	----------	235.00 µl
1M Tris-HCl (pH 8.0)	----------	188.00 µl
SDS (20%)	----------	30.00 µl
Proteinase K	----------	30.00 µl

Make up the volume to 1000 µl autoclave and store at room temperature.

g) Ethanol 70%

Ethanol 99.9 %	----------	70.00 ml
Distilled Water	----------	30.00 ml

h) Agarose

	0.7% gel conc.	**1% gel conc.**
Agarose	0.28 g	0.6 g
TAE (1X)	40.0 ml	40.0 ml

Agarose was added in 1X TAE buffer and dissolved by heating.

i) Ethidium Bromide

Ethidium Bromide	----------	10.00 mg.

Dissolved in 1 ml autoclaved distilled water and stored at room temperature.

j) TAE Buffer (50X) stock solution

Tris base	----------	242 g
Glacial Acetic Acid	----------	57.1 g
EDTA (0.5M; pH 8.0)	----------	100.0 ml

Double distilled water was added to a volume of 1000ml, filtered and autoclaved.

PROCEDURE:

Cell lysis:

1. 20-30 mg fish muscle was taken in 2 ml vial and mixed with 200 µl of 940 µl of lysis buffer.

2. Muscle was crushed with micropistel, after complete crushing of tissue rest 740 µl of lysis buffer was added.

3. 30 µl of 20% SDS and 15 µl of Proteinase K (stock solution 10 mg/ml) was added and incubated at 48 ºC for 2 to 3 h in a water bath till complete lysis occurs (the solution should look clear).

Phenol Extraction:

1. Equal volume of Phenol (Tris saturated, pH 8.0): Chloroform: Isoamyl alcohol (25:24:1) was added to the lysed cell mixture and mixed to form a uniform Suspension by gently inverting the tube 40-45 times.

2. The mixture was centrifuged at 12,000 rpm for 12 min at room temperature.

3. The upper aqueous phase was gently aspirated using a wide bore sterile micropipette, without disturbing the interphase containing protein and then transferred to a fresh micro tube (2 ml).

4. to the aqueous layer, equal volume of phenol: chloroform (4:1) was added and mixed properly.

5. Now the mixture is centrifuged at 12,000 rpm for 12 min at room temperature.

6. The upper aqueous phase was transferred to a fresh micro tube and equal volume of chloroform was added to the aqueous phase and mixed gently.

7. The mixture was centrifuged at 10,000 rpm for 5 minutes at room temperature.

8. The upper aqueous phase was transferred to a fresh 1.5 ml micro tube for DNA precipitation.

DNA precipitation:

1. For DNA precipitation, equal volume of Isopropyl alcohol and 0.2 volumes of 10 M Ammonium acetate was added to the aqueous layer and mixed gently for DNA precipitation. The micro tubes were kept at -20 ^{0}C for 30 min.

2. After precipitation, the mixture was centrifuged at 12,000 rpm for 12 min at room temperature to get the DNA pellet.

3. The supernatant was discarded and the DNA pellet was washed with 500 µl of chilled 70% ethanol for desalting. For this, 70% ethanol was added to micro tube, mixed well and centrifuged at 10,000 rpm for 5 min. This step was repeated for two times for optimum desalting.

4. The DNA pellet was dried by keeping the micro tube (with lid open) in a sterile incubator at 37 ºC overnight.

5. After drying, the DNA was dissolved in 50 µl of sterile water. The lid of the micro tube was sealed with Parafilm wax and kept in water bath at 65 ºC for 30 min.

6. The DNA can be stored at -20 ºC for downstream applications.

7. Before Agarose gel electrophoresis we kept the sample in 50 µl Milli Q water at 37°C at room temperature

QUANTIFICATION OF DNA BY AGAROSE GEL ELECTROPHORESIS:

Agarose gel electrophoresis is a method used in biochemistry and molecular biology to separate, identify and isolate DNA, or RNA molecules by size. This is achieved by moving negatively charged nucleic acid molecules through an Agarose matrix in an electric field. Shorter molecules move faster and migrate farther than longer ones. The most common dye used to make DNA or RNA bands visible for Agarose gel electrophoresis is Ethidium Bromide, usually abbreviated as EtBr. It fluoresces under UV light when intercalated into DNA (or RNA). By running DNA through an EtBr-treated gel and visualizing it with UV light, distinct bands of DNA become visible. EtBr is a potent carcinogen, however, and safer alternatives are available.

This is the standard method to separate, identify and purify DNA fragments. This technique separates DNA molecule according to their molecular weights in the same manner as a sieve. The technique is simple, rapid to perform and capable of resolving DNA fragments that cannot be done adequately by the other procedures such as density gradient centrifugation.

1. After taking out sample from Mile Q water 10 mg Dye (6X) + 10 ml NaOH (10 Mm) + 5 ml Glycerol (50%) + 9400 µl DDW was added to it.

2. 0.20 g of Agarose (1%) was dissolved in 20 ml TAE buffer in a 100 ml conical flask and was boiled to dissolve Agarose completely.

3. Now the Agarose was cooled at room temperature and 2µl EtBr was added to it.

4. Then it is poured and allowed to solidify at room temperature.

After the gel solidifies, the DNA samples along with standard DNA marker were loaded. The electrophoresis was carried out at a constant voltage of 55V. After 30 minute, the gel was observed under ultraviolet light of UV trans-illuminator and the image was captured & analyzed.

QUANTITY AND QUALITY ANALYSIS OF DNA:

The quantity and quality of DNA was determined by following ways:

- using spectrophotometer
- using U.V. transilluminator
- using gel documentation system

For spectrometric quantization of DNA, optical density was recorded at 260 nm and 280 nm wavelengths. The reading at 260 nm allows calculation of the concentration of nucleic acid in the sample. An optical density of 1 corresponds to approximately 50 ng/ µl of double-stranded oligonucleotide. The standard conversions of spectrophotometeric reading i.e., OD to nucleic acid concentration are as follows:

1 A260 unit of double-stranded DNA = 50 ng / µl

1 A260 unit of single-stranded DNA = 33 ng / µl

The concentration of unknown double standard DNA sample is estimated using the following formula:

$$\textbf{Concentration (µg/µl)} = \frac{OD_{260} \times \text{DILUTION FACTOR} \times 50}{1000}$$

Gel electrophoresis is the standard method to separate, identify and purify DNA fragments. This technique separates DNA molecules according to their molecular weights in the same manner as a sieve. Bands containing as little as 20 ng of the DNA can be detected by directly examining the gel. The DNA was run on 1% Agarose gel for 1 hour, the gel was observed under U.V. transilluminator and the DNA quality was determined by observing the gel.

In gel electrophoretic determination, the DNA was loaded in the gel along with DNA marker. The gel was analyzed on Syngene Gel Documentation system and DNA concentration was determined as per the manual.

MAJOR STEPS IN ISOLATION OF GENOMIC DNA FOR PCR AMPLIFICATION

Fig 6: Collection of tissue by sacrificing the fish

Fig 7:Add lysis buffer in collected tissue sample

Fig 8:Tissue sample was put in water bath for complete lysis

Fig 9: Centrifuge was taken to separate chemicals and cell extracts

Fig 10: Various chemicals like phenol-chloroform and isoamyl alcohol were Added with repetitive centrifuge

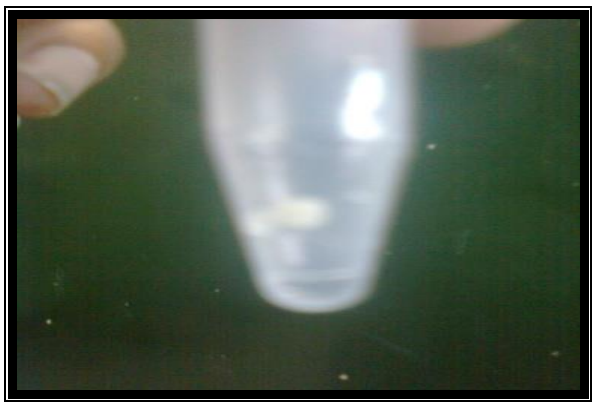


Fig 11: DNA precipitated as pallet which was carefully dried and dissolved

Fig 12: Routine use agarose was weighted in weight balance

Fig 13: Gel was molted ethidium bromide added and casted in a casting tray having combs

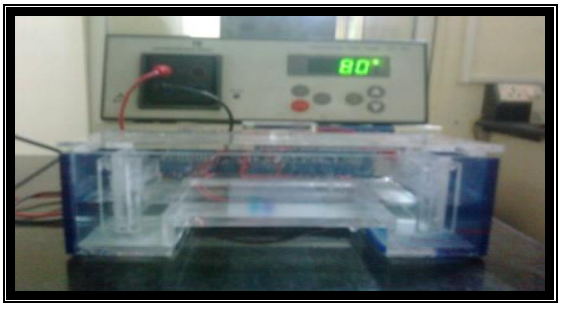


Fig 14: Isolated DNA samples were loaded in wells with bromo phenol blue dye

& Electrophoresis was carried out

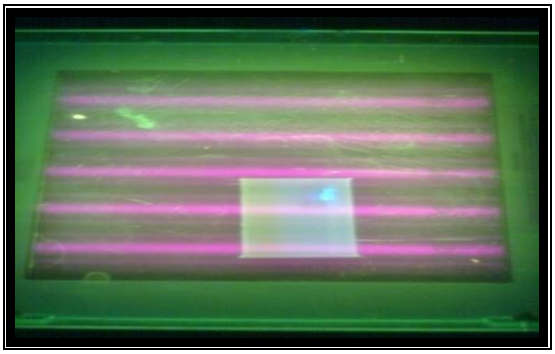


Fig 15: Gel slabs was observed on U.V transilluminator to score DNA appeared as Flourishing bands

Fig 16: Gel documentation system was used to analysed the gel slab

Fig 17: Concentration of isolated DNA was measured by U.V/ Visible Spectrophotometer

Fig 18: PCR amplification of isolated DNA was done by using specific primers

PCR products were analysed by Agarose Gel Electrophoresis and

Send for sequencing

Fig 19: Isolated DNA, Amplicons and PCR reagents were stored in -20

Cryopreservation unit.

PCR REACTION:

The fish DNA was amplified in order to study different ribosomal regions.

Reaction Mixture setup:

The reaction mixture for 50µl volume was setup and above reaction mixture was mixed gently. The components of the reaction mixture were as follows:

S.No	STOCK REAGENTS	FINAL CONC.	VOLUME/REACTION
1	Double distilled Water	_____	37.0µl
2	10x Polymerase Buffer	1X	5.0µl
3	10mM dNTPs	0.2mM	1.0µl
4	25 mM $MgCl_2$	2.0mM	2.0 µl
5	10pmol/µl Forward Primer	10pmol	1.0µl
6	10pmol/µl Reverse Primer	10pmol	1.0µl
7	1U/µl Taq DNA Polymerase	1U	1.0µl
8	50ng Genomic DNA	_____	2.0µl

Table 1: Showing the components required for PCR.

Amplification of target sequence from genomic DNA extracted from muscle may be carried out using suitable primers. The PCR reaction may contain 20-25ng genomic DNA, 2.5µl of 10X PCR buffer, DDW (as per the reaction volume), 2.0mM $MgCl_2$, 0.5µl of 10mM $dNTP_s$ mix, 5-10pmol of each forward and reverse primers and 0.5U Taq Polymerase.

PCR Amplification:

The amplification is carried out in BIORAD DNA ENGINE Thermo Cycler.

The amplification was performed in a total volume of 50µl containing 30-50ng of template DNA.

The PCR conditions are usually initial denaturation at 94°C for 3minutes followed by 32cycles of denaturation at 94°C for 1minutes, annealing at 30-60°C(depending upon primer size) for 30seconds, primer extension at 72°C for 1minute followed by final extension at 72°C for 10minutes. The amplified product (8µl) may be used to load in 1.5% Agarose for its analysis for 1hour. The gel was taken out and placed on UV transilluminator to visualize the amplified products. Then the photograph was captured using the Genesnap software and the Gel Documentation System.

PCR PROGRAM:

S.No.	Step	Time	Temperature
1	initial denaturation	3 minutes	94°C
2	Denaturation	30 seconds	94°C
3	Annealing	30 seconds	55°C
4	primary extension	1 minutes	72°C
Start of repeat for 32 cycles			
5	final extension	10 minutes	72°C

Table 2: Showing the PCR reaction

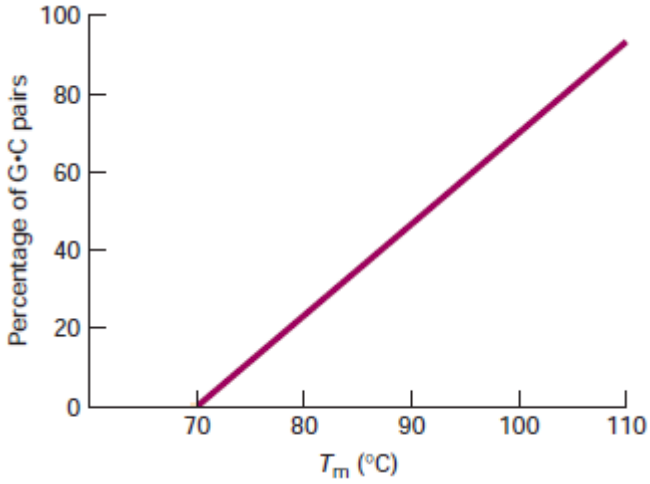
Primers:

Combinations of primers were designed to amplify contiguous and overlapping fragments of rDNA gene. The primer designing was based on highly conserved regions of an alignment having sequences from several eukaryotes

Formula for Melting temperature:

Melting temperature for polynucleotide sequence can be calculated by formula given as below

$$T_m = 4 \times (\text{Number of GC Base Pair}) + 2 \times (\text{Number of AT Base Pairs})$$



bases in a double-stranded DNA sample have denatured is denoted T_m (for temperature of melting). Light absorption by single-stranded DNA changes much less as the temperature is increased. (b) The T_m is a function of the G·C content of the DNA; the higher the G+C percentage, the greater the T_m.

Primer details

Primers of ribosomal ITS region were designed with the help of different primer designing software like PRIMER3. There are also two sets of forward and reverse primers were designed:

ITS 2 Forward primer (5'-GCAGGACACATTGATCATCGACAC-3')

Tm (melting temperature)	**72**
Number of bases	24
GC%	50
AT%	50

Table 3: Description of ITS 2 forward primer

ITS 2 Reverse primer (5'-GGCTCTTCCCTCTTCGCTCGC-3')

Tm (melting temperature)	**70**
Number of bases	21
GC%	66.6
AT%	33.4

Table 4: Description of ITS 2 reverse primer

ITS (for internal transcribed spacer) refers to a piece of non-functional RNA situated between structural ribosomal RNAs (rRNA) on a common precursor transcript. Read from 5' to 3' this polycistronic rRNA precursor transcript contains the 5' external transcribed sequence (5' ETS), 18S rRNA, ITS1, 5.8S rRNA, ITS2, 28S rRNA and finally the 3'ETS. During rRNA maturation, ETS and ITS pieces are excised and as **non-functional maturation by-products** rapidly degraded. Genes encoding ribosomal RNA and spacers occur in tandem repeats that are thousands of copies long, each separated by regions of non-transcribed DNA termed *intergenic spacer* (IGS) or *non-transcribed spacer* (NTS). Sequence comparison of the ITS region is widely used in taxonomy and molecular phylogeny because it a) is (due to the high copy number of rRNA genes) easy to amplify even from small quantities of DNA, and b) has a high degree of variation even between closely related species. This can be explained by the relatively low evolutionary pressure acting on such non-functional sequences.

The ITS region is now perhaps the most widely sequenced DNA region in fungi (Peay et al., 2008). It has typically been most useful for molecular systematic at the species level, and even within species (e.g., to identify geographic races). Because of its higher degree of variation than other genic regions of rDNA (for small- and large-subunit rRNA), variation among individual rDNA repeats can sometimes be observed within both the ITS and IGS regions. In addition to the standard ITS1+ITS4 primers used by most labs, several taxon-specific primers have been described that allow selective amplification of fungal sequences (e.g., see Gardes & Bruns 1993 paper describing amplification of basidiomycete ITS sequences from mycorrhiza samples).

RESULT

Genomic DNA

The phenol – chloroform method was used for genomic DNA isolation, good quality DNA was isolated from 40mg muscle sample. The genomic DNA concentration ranged between 40-80 ng/µl with an average concentration of 50ng/µl from specimen of *P.nigrus,* P.flavomaculatus *and L.parsia* RNAase treatment was not given to isolate the genomic DNA, hence contamination of genomic DNA was also observed in the DNA samples. However, the genomic DNA isolated from muscles was free from other impurities like proteins, and phenols etc.

DNA of three different species of marine fishes (*P.nigrus, P.flavomaculatus and L.parsia*) was isolated from tissue samples.

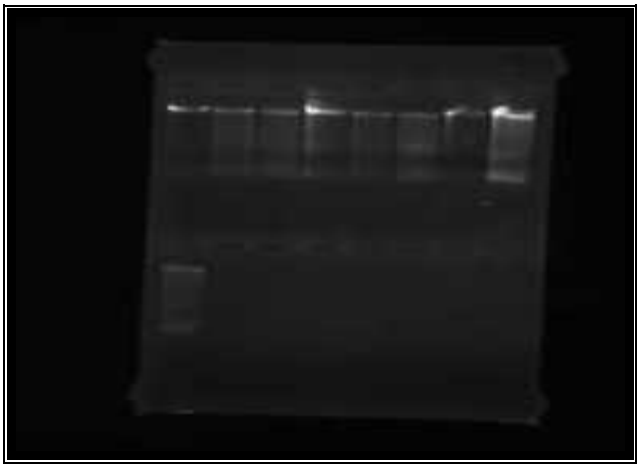


Fig 20: Isolated DNA

Primers Used:

ITS 2 F1: 5’-GCAGGACACATTGATCATCGACAC-3’

ITS 2 R1: 5’-GGCTCTTCCCTCTTCGCTCGC-3’

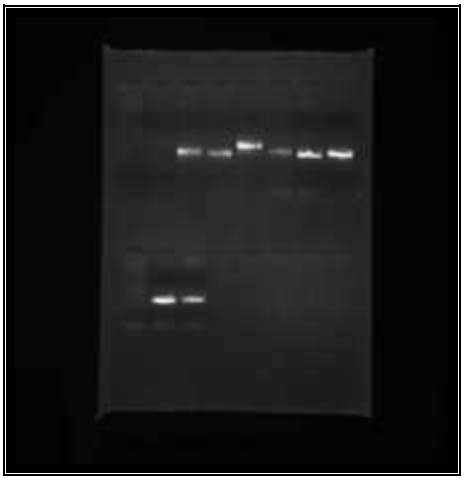


Fig 21: Isolated DNA

PCR product

Forward and Reverse primers of ITS 2 ribosomal DNA was employed for amplification of specific regions of ITS 2 ribosomal DNA from nuclear genomic region of different freshwater catfish species.

The designed sets of primers consistently gave by PCR. The amplified fragments covered the coding regions of ITS 2 ribosomal gene of the undertaken species. The sequence of these primers is conserved among the studied species. Same sizes of PCR products in the undertaken species were observed which was located in the variable expansion segments that were subsequently excluded from phylogenetic analysis.

The amplified region ITS 2 ribosomal gene was obtained from polymerase chain reaction with the help of isolated DNA and suitable ITS 2 ribosomal primers.

SEQUENCES ANALYSIS

AMPLICONS SEQUENCES

Plectorinchus flavomaculatus (ITS 2)

CAGGACACCATTGATCATCGACACTTCGAACGCACCTTGCGGCCCCGGGTTCCTCCCGGGGCT
ACGCCTGTTTGAGGGTCGCTTTGCCATCAATCGGAAACGCCCGCGTTTCCGCGGCTGGGGCAG
TCGCAGGCGAGCCTTCGCCTTTGTCCCCCCAAGTGCAGACTCAGTGTGCTGGGACGCTCGGCG
CGGTGCACCGTGGGCCTGTCCGGGGCCCACCCTGCCTTCCTTTCCGTTGGCTTTTCACACCTCT
CTCTCCCTCCTTTTCCGGCCTCCGGGCCGGGGAGGGGCGCCCTGTCGGGGCCCGCATGAGTCG
GGCGCGGCTGCCGGTGGATGACTCCGCGTTTTTTCCGTGCTGACCGCGTTACGCGTGCGCCGGT
TTCCGTGCGGGGATCGCTCTAGCGGGGGTTTGAGGGGTTAGCCTTCGGGCGGGGTCCGGCACG
GGTGCGGACCACCGGCTTGCCCCTCTCCCGCGTCGGGTTACTCGGCACGCGCCTTTTTTTCCCC
CCCCCCGGAGCTCGTCGTGAGCCTCCTGTCAACCGGCCGGCTGGCGGAGCCACGCACCCAGAT
GCATTCGACTACGACCTCAGATCAGACGAGACAACCCGCTGAATTTAAGCATATTACTAAGCG
GAGGAAAAGAAACTAACCAGGATTCCCTCAGTAGCGGCGAGCGAAGAGGGGAAGAGCCA

Plectorhinchus nigrus (ITS2)

GCAGGACACATTGATCATCGACACTTCGAACGCACCTTGCGGCCCCGGGTTCCTCCCGGGGCT
ACGCCTGTCTGAGGGTCGCTTTGCCATCAATCAGAAACGCCCGCGTTTCCGCGGCTGGGGCAG
TCGCAGGCGACCTCCCGGTCGCCTTCGTCCCCCCAAGTGCAGACCTGTCCGGGACGCCCGGCG
CGGTGCACCGTGGGCCTGCCCGGCCCACCCTGCCTTCCTCTCCGTTGGCTTCTCCCGCCTCTCT
CTCCCTCCTTCTCCGGCCCCCGGGCCGGGGAGGGGCGCCCTGTCGGGCCCGCATGAGTCGGGC
GCGGCTGCCGGTGGACGACTCCAAGTCTCCGTGCTGACCGCGTTACGCGTGCGCCGGTTTCCG
TGCGGGGACGCTCTAGCGGGGGTCAGAGGGGTTAGCCTTCGCGGGGGTCGGCACGGGTGCGG
ACCCGGCTTGCCCCTCTCCCGCGTCGGGGCACCCGGCGCTCTTCTCCGGAGCTCGCCGTGAGCC
TCCGCCCCTGGCCGGCGGAGCCACGCACACACGCCGTTCGACTACGACCTCAGATCAGACGAG
ACAACCCGCTGAATTTAAGCATATTACTAAGCGGAGGAAAAGAAACTAACCAGGATTCCCTCA
GTAGCGGCGAGCGAAGAGGGA

***Liza Parsia (*ITS 2)**

ACATTGATCATCGACACTTCGAACGCACCTTGCGGCCCCGGGTTCCTCCCGGGGCTACGCCTGT

CTGAGGGTCGCTTTGCCATCAATCGGGAGGCCTCGGCCTCCCGCGGCTGGGGGCTTTTCGCAG

GCGGCCTCGGCCGTCTTCGTCCCCCTAAGTGCAGACCTTTCGCTTCTCTCCCTCTTTTTTGGGCC

CCCTTCCTCGGCTCCCTTCCCCCTACCTCTCTCCCTTCCACCGCCTCCGGGGGGTTGGTTGTGCG

TCAGGCTAGGGGGGCGCCTTCCCCGATCGAGCTCCGCACGCTGCCACGGGGCGCGGCTGCCGG

TGGACGACACTCGTCTCCGCGCTGACCGCGTCTCCCCACCGCGTGCGTGGCGCTCAAGGGGGG

TCTGCGTCGGGTTCCGAGAGGGGCGGGGTTCCGAGAGGAAGAAGTCGAGAGACACTCCAATC

ACCAGCCAAAAACGTGAGCCTCCGTGCGTCGGAGCCACGGTTCAGCCTTTTCGACTACGACCT

CAGATCAGACGAGACGACCCGCTGAATTTAAGCATATTACTAAGCGGAGGAAAAGAAACTAA

CAAGGATTCCCTCAGTAGCGGCGAGCGAAGAGGGAAGAGCCAAGAGCC

Sequences of all three fishes on the basis of which we will find out the phylogenetic relationship between the fishes using the software MEGA 4.

Construction of Phylogenetic tree:

Phylogenetic relationships are often represented graphically, either in form of phylogenetic trees (often unrooted), in which evolutionary distance is measured in the term of horizontal branch length, or as dendograms, in which evolutionary distance is measured along the length of line segments. Dendogram of taken species was constructed by using MEGA 4 software.

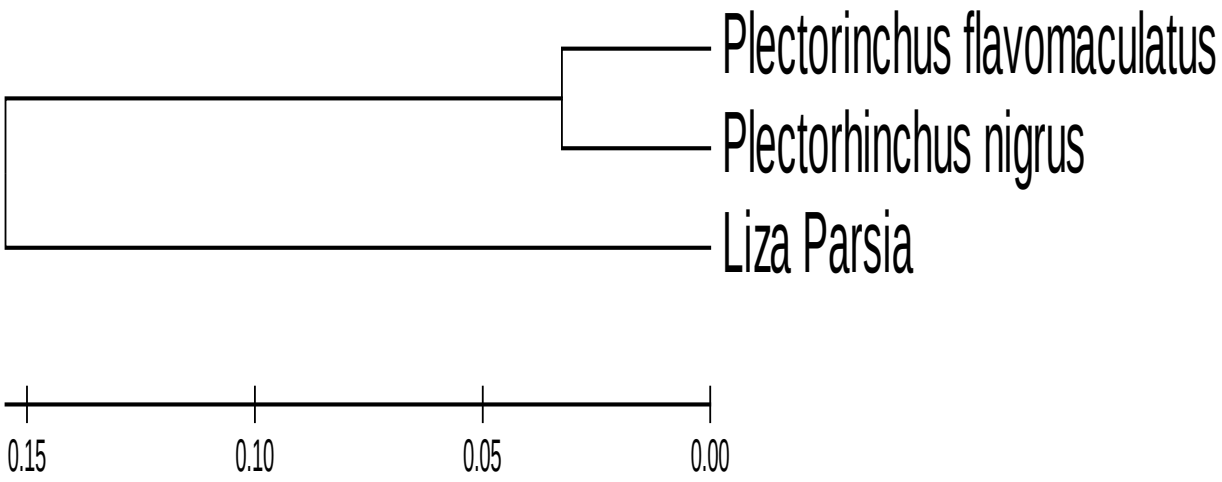


Fig.22: Tree showing phylogenetic relationship among the species studied using Neighbour-Joining method

The phylogeny among three species was constructed using MEGA 4 software in which *Plectorhinchus flavomaculatus*, *Plectorhinchus nigrus* were clustered in one cluster while *Liza Persia* is separated in another cluster.

PHYLOGENETIC DETAILS OF SEQUENCES

	T	C	A	G
Plectorinchus flavomaculatus	20.4	33.8	14.5	31.4
Liza Parsia	19.8	33.8	16.7	29.7
Plectorhinchus nigrus	17.2	36.6	14.4	31.8
Avg.	19.1	34.8	15.1	31.0

Table 5: It shows the GC and AT content of particular as well as over all average of the content of all three fishes i.e

GC content = G+C = 92.9+104.2= 197.1/3=65.7

AT content = A+T = 45.6+57.4 = 103/3=34.3

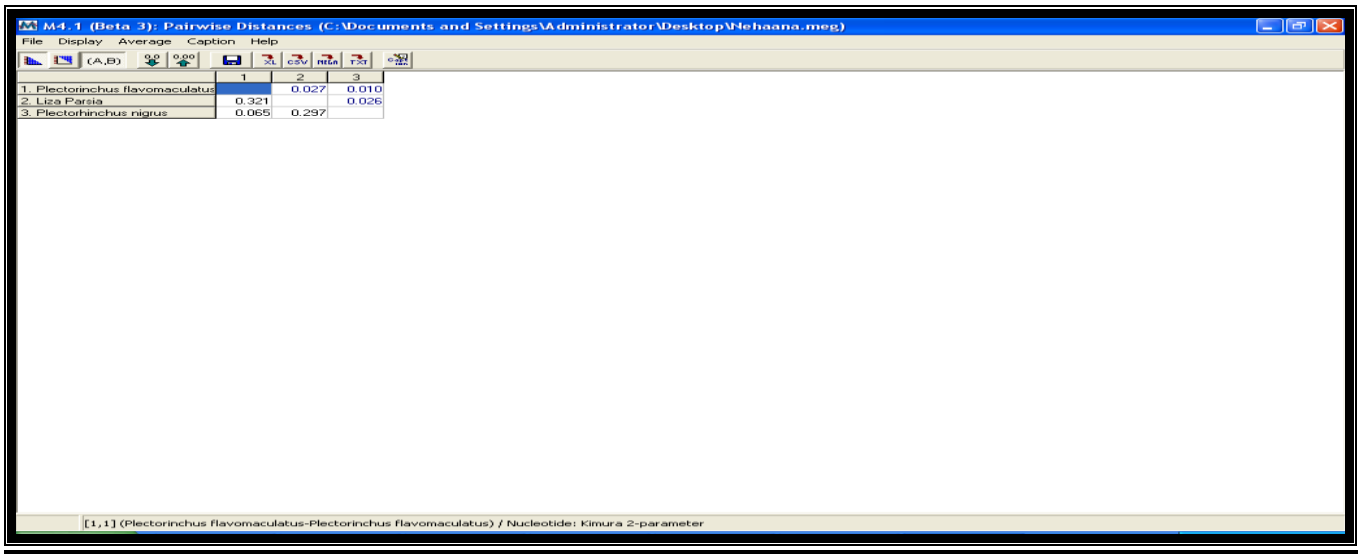


Table 6: Showing distance (below diagonal) and standard error (above diagonal)

The minimum distance is seen between *Plectorhinchus flavomaculatus* and *Plectorhinchus nigrus* i.e. 0.065 and the maximum distance is seen between *Plectorhinchus flavomaculatus* and *Liza parsia* i.e. 0.321 and the distance between *Liza parsia* and *Plectorhinchus nigrus is 0.297.*

Variable sites	**170**
Conserved sites	**489**
Total base	**706**

Table7: showing conserved and variable regions

After analysis, total sites obtained were 706 out of which 489 were conserved and 170 were variable among the three fishes.

DISCUSSION

For molecular characterization, live specimens of *Plectorhinchus nigar, Plectorhinchus flavomaculatus* and *Liza parsia* were provided by NBFGR repository. The study was performed with the objective for molecular characterization of the fishes. Although genetic method are very effective but unfortunately it cannot appear in vogue for fisheries, this is because perhaps currently available methods are time taking, laborious, and expensive for applying on thousand of samples **(Palumbi and Cipriano., 1998).**

DNA based analysis required high quality DNA samples for genetic analysis DNA of high quality is best for screening of large samples at molecular level.

In this project DNA extraction was done by using PHENOL – CHLOROFORM – ISOAMYL ALCOHOL (PCI) differential solubility method **(Russell and Sambrook, 2001).** Few modifications in this method for high quality and concentration of DNA were observed. Lysis buffer, SDS and Proteinase K are used for tissue digestion**.** A better dissociation of the tissue was obtained when the digestion was done at 48°C **(Wasko *et al.*, 2003)**.

I was careful for phenol treatment because delay in exposure to phenol causes sheared and low quality DNA. In most cases morphologically similar species can announced distant by applying the global technique of gene amplification, the **PCR** in which ITS 2 ribosomal gene was targeted for amplification in a rapid and reliable method. PCR allowed to study of genetic diversity among the species throughout the evolution.

Important nessicities of a genetic assay desired routine analyses and accuracy, speed and reliability. Various key designs are considered along with these properties.

There was specific and estimated temperature for primers under high accuracy. Success of PCR amplification was found very sensitive with various qualitative factors which includes, Taq polymerase inhibiting factors. Amplification reactions could be effected by dilution errors and loss of aqueous substance by high temperature. To overcome this I had used wax wrapping. **(Sandy Primrose, Richard Twyman and Bob Old., 2003)**

Complete reaction failure are not uncommon for PCR reaction, to carry the reaction in a desired control temperature adjustment is very important it affects primer annealing and extinction. Concentration of reaction mixture (master mix) component e.g. $Mgcl_{2,}$ dNTPs, and Taq enzyme ensures reliability of amplification.

After complete PCR cycle **Agarose gel electrophoresis** thought to best method for imaging rough score about amplicon size. During Electrophoresis buffers were used e.g. TAE, TBE. Chemical staining properties of DNA was used to trace its position into the gel slab by under exposure of U.V. light. Images could be acquired with using gel doc system.

Molecular evolution of under taken species could be traced after sequence analyses of amplicons and multiple sequence alignment. Analyses showed that through molecular evolution some regions of ITS 2 RNA gene were became conserved as Consensus but some regions were greatly variable.

Phylogenetic hypotheses of the evolutionary relationship among species provides framework for compatible research or mechanism of diversification and speciation. These phylogenetic data will also important for conservation strategy.

This is a mean of quantifying evolutionary distinctiveness and similarities to resolve taxonomic ambiguities involving rare texa. My study demonstrates that partial ITS 2 Ribosomal gene provides strong phylogenetic relationship between all the subject species belong to the different genus.

PRECAUTIONS IN HANDLING OF DNA

1- In General Enzymes Are Added In Last.

2- For best results fresh blood stored at room temperature up to 48 hrs. Should be used; however blood stored at -20^0 c. can be used up to 30 days; however yield will reduce about 10-15%.

3- Blood can be stored in ethanol (10.25ml blood +1.25 ml of 95% ethanol) for months vigorous pipetting should be avoided.

4- Tips with wider orifice used for genomic DNA isolation.

5- Genomic DNA should not be vortexed.

6- DNA should be stored in aliquots at 4^0C and -8^0C for medium to long term storage.

7- Genomic DNA should be air dried after ethanol precipitation and over drying of DNA should be avoided for plasmid and other smaller DNA can be air / vacuumed dried.

SUMMARY AND CONCLUSION

Molecular characterization of partial ITS 2 ribosomal gene in fresh water catfish species *P.nigrus*, *P.flavomaculatus* and *L.parsia* was done. DNA extraction was done by applying phenol - chloroform - isoamyl alcohol method from muscle sample. Good quality of genomic DNA was obtained with an average of 50 ng/µl. Extracted DNA samples were used as template to amplify 18s ribosomal region using three different primers with two sets of forward and reverse, which generates amplicons of 395 to 692 bp in all three species. Sequencing of these regions was outsourced on payment basis. Analysis of the sequence performed using CLUSTAL-W multiple sequence alignment and BIOEDIT. The BLAST was carried out for both species and phylogenetic tree was constructed on account of the molecular evolution and conservation of this region.

Evolutionary biology and systematic both have huge importance of comparative DNA analyses. large amount of comparative sequence data will be required to answer most molecular systematic questions, but this will be a labour intensive effort and one of the toughest task about systematic.

Presence of long conserved regions in ITS 2 ribosomal gene indicates the inter species similarities in this region. Although the sequence length found species specific with slight variation and use to study long term evolution. These sequences can be used for taxonomic identification.

Bioinformatic tools for determining phylogeny

Along with laboratory approach computational procedures were also used to analyze the sequences. Computer programs and software were used as given below:

1. CLUSTAL W- EBI this was used for multiple sequence alignment
2. BIOEDIT was used to read and edit sequencing report obtained as Electroferogram.
3. PRIMER 3 is one of the best on line primer synthesizing application.
4. OLIGO CALC was used to calculate thermo physical and molecular properties of any polynucleotide sequence.
5. MEGA 4 was used to discover phylogenetic relationships and construct Dendograms. It can calculate phylogenetic distances and explore evolutionary relationship.

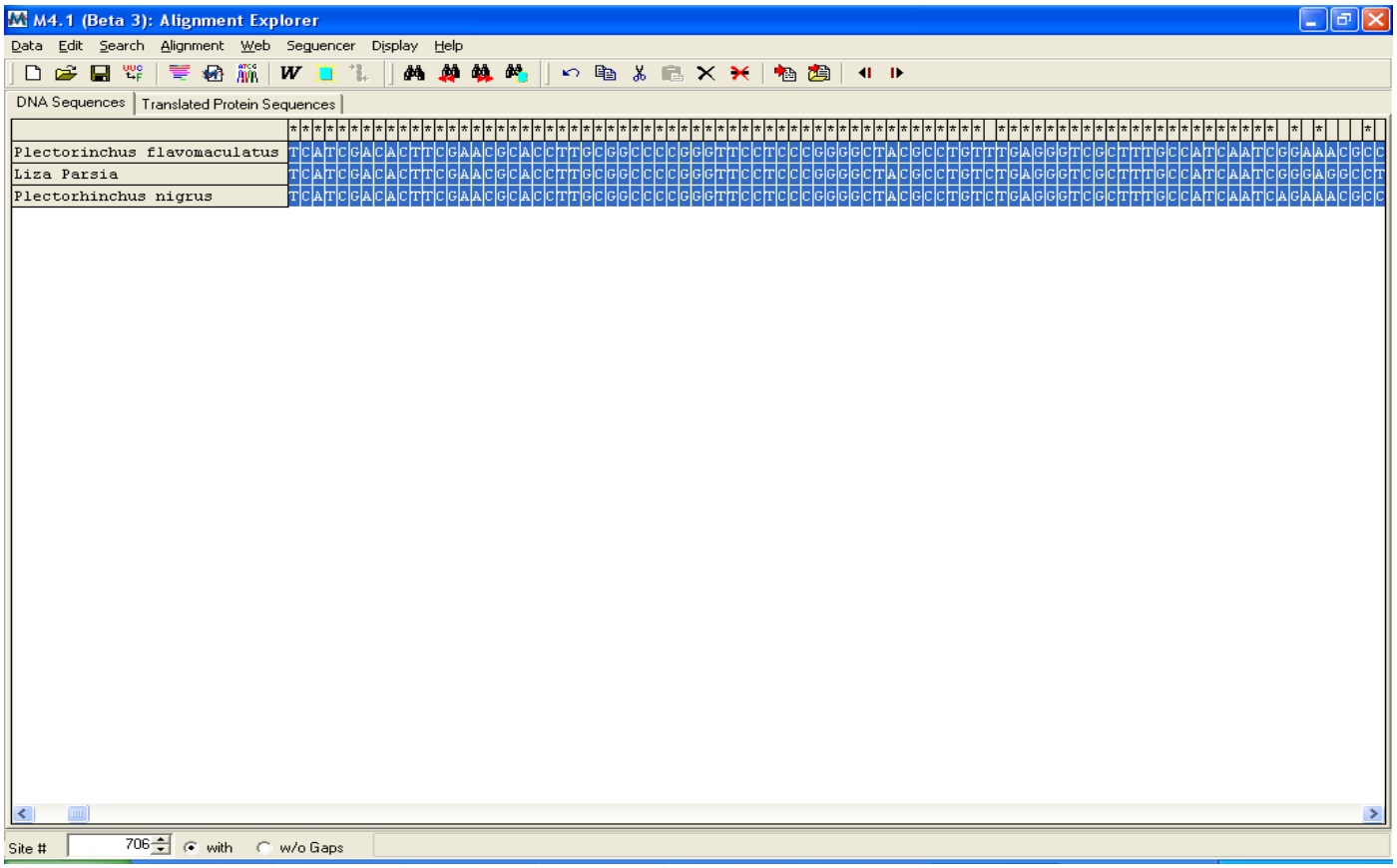


Fig 23: Application window of MEGA 4 software

Appendix

Reagents required for DNA isolation

Stock solutions:

1. **0.5 M Tris-Cl (pH 8.0)**

   Tris base - 3.028 g

   Distilled water - 40.0 ml

   The pH was adjusted to 8.0 using 1N HCL

   Final volume was makeup to 50ml

   Solution was Stored at 4^0 C.

2. **0.5 M Tris Cl (pH 8.3)**

   Tris base - 3.028 g

   Distilled water - 40 ml

   The pH was adjusted to 8.0 using 1N HCL

   Final volume was makeup to 50ml

   Solution was Stored at 4^0 C.

3. 0.5 M EDTA (pH 8.0)

Na_2 EDTA.$2H_2O$ - 9.34g

Distilled water - 40 ml

The pH was adjusted to 8.0 using 0.5 M NaOH.

Final volume was makeup to 50ml

Solution was Stored at 4^0 C.

4. 0.5 M Tris Cl (pH 7.5)

Tris base - 3.02 g

Distilled water - 40ml

The pH was adjusted to 7.5 using 1N HCL

Final volume was makeup to 50ml

Solution was Stored at 4^0 C.

5. 5X TAE

Tris base - 12.10g

0.5 m Na_2EDTA.$2H_2O$ (pH 8.0) - 5.0 ml

Glacial acetic acid – 2.85 ml

Final volume was makeup to 50ml

Solution was Stored at 4^0 C.

6. **Bromo Phenol Blue dye**

   Bromo phenol Blue - 2.5mg

   Sucrose - 40.0 mg

   Dissolved in 1 ml distilled water

   Solution was autoclaved and stored at 4^0 C.

Working solutions:

1. **High TE**

   Stock 0.5 M Tris-Cl (pH 8.0) - 20ml

   Stock 0.5M $Na_2EDTA.2H_2O$ (pH 8.0) - 8.0 ml

   Final volume was makeup to 50ml

   Solution was autoclaved and stored at 4^0 C.

2. **Incubation buffer**

   0.5M Tris Cl (pH 8.0) – 2ml

   0.5M EDTA – 2 ml

   Distilled water - 96.0 ml

   Solution was autoclaved and stored at 4^0 C.

3. Proteinase K

Proteinase K- 10 mg

Autoclaved distilled water-500 µl

Proteinase K was dissolved into above volume of distilled water.

Solution was Stored at -20^0 C.

4. RNAase Buffer

0.5M Tris Cl (pH 7.5) – 0.2 ml

NaCl (0.292g in 10 ml) - 0.3 ml

Distilled water – 9.5 m

Solution was autoclaved and stored at 4^0 C.

5. RNAase

RNAase - 10 mg

Autoclaved RNAase buffer- 1 ml

RNAase was dissolved in RNAase buffer with in a tube.

Tube was kept in boiling water for 15 minutes.

Stored at - 20^0 C.

6. 3 M Sodium Acetate (pH 5.2)

Sodium acetate- 12.4 g

Distilled water- 20ml

The pH was adjusted to 5.2 using glacial acetic acid

Final volume was makeup to 50 ml

Solution was Autoclaved and stored at 4^0 C.

7. TE Buffer

Stock 0.5 M Tris Cl (pH 8.0) – 2.0ml

Stock 0.5 M $Na_2EDTA.2H_2O$ (pH 8.0) – 0.02 ml

Final volume was makeup to 100 ml.

Solution was autoclaved and stored at 4^0 C.

8. 0.5 X TAE electrophoresis buffer

5X stock TAE – 25 ml

Distilled water- 225 ml

Procedure of Phenol Saturation Tris - HCL (pH 8.0)

Required chemicals

Water saturated phenol - 500 ml

0.5M Tris HCl- (pH 8.0) - 1000 ml (60.56g of Tris base in 1000ml water)

0.1 M Tris HCl (pH8.0) - 1500 ml (For 300 ml of 0.5M Tris-HCL adds 1200 ml of dis.water)

Steps:

1- 0.1% to 8%Hydroxyquinoline was added to water saturated phenol.

2- Flask containing above solution was wrapped with aluminium foil to avoid light reaction.

3- 500 ml of 0.5 M Tris HCL was added.

4- Solution was stirred using magnetic stirrer for 15 minutes.

5- Solution was kept for 30 minutes to settle down phenol.

6- Supernatant was decanted.

7- 500 ml of 0.1 M Tris – HCL was added.

8- Steps of stirring, settling, and decanting were repeated twice with 0.1M Tris HCL.

9- pH of decanted supernatant was checked and finally find 8.0.

10- 500 ml of 0.1 M Tris - HCl was added to phenol.

11- Stored in dark bottles and covered with aluminium foil.

LIST OF ABBREVIATIONS

%	-	Percentage
°c	-	Degree centigrade
µg	-	Microgram
ATP	-	Adenosine-5-triphosphate
bp	-	Base pairs
cm	-	Centimetre
dATP	-	2-deoxyadenosine-5-triphosphate
ddNTP	-	Di deoxy nucleotide triphosphate
DDW	-	Double distilled water
dNTP	-	Deoxy nucleotide triphosphate
EtBr	-	Ethidium Bromide
Fig	-	Figure
g	-	Gram
mg	-	Milligram
min	-	Minutes
ml	-	Millilitre
ng	-	Nanogram
OD	-	Optical Density

PCR	-	Polymerase Chain Reaction
rDNA	-	Ribosomal DNA
Rpm	-	Revolutions per minute
SDDM	-	Sterile Double Distilled Water
SDS	-	Sodium Dodecyl Sulphate
TAE	-	Tris Acetic Acid
Taq	-	Taq Polymerase
TBE	-	Tris Borate EDTA
TE	-	Tris EDTA
W/Vol	-	Weight/Volume

REFERENCES

A. Begg G.A., Keenan C.P., and Selin M.J. (1998). Genetic Variation and Stock Structure of School Mackerel and Spotted Mackerel in Northern Australian Water. *J. Fish Biology* **53**: 543-559.

B. Bernardi G. and Talley D., (2000). Genetic Evidences for Limited Dispersal in the Costal California Killifish, Fundulus parvipinnis. *J. Exp. Mar. Bio. Ecol.* **255**: 187-199.

C. Bruce Albert, (2004). Molecular Biology of Cell: Garland Science.

D. Felsenstein J., (1985). Confidence Limits on Phylogenies: an approach using the bootstrap. *Evol*. **39**: 783-791.

E. Field K. G., Olsen G. J., Lane D. J., Giovannoni S. J., Ghiselin M. T., Raff E. C., Pace N. R. & Raff R. A., (1988). Molecular Phylogeny of the Animal Kingdom. *Science.* **239**: 748-753.

F. Haldane J. B. S., (1955). Population Genetics. *J. New Biology.* **18**: 34–51.

G. Hamilton, W.D., (1964). The Genetical Evolution of Social Behaviour. *Journal of Theoretical Biology.* **7**: 1-52.

H. Higgins D. G., Bleaby A .J. and Fucsh R., (1992). CLUSTAL W: Improved Software for Multiple Sequences Compute. *Appl. Biosci*. **8**: 189-191.

I. Hillis D. M., Dixon M.T., (1991). Ribosomal DNA: Molecular Evolution and Phylogenetic Inference.The quarterly review of biology. **66**: 411-453.

J. Kimura M., (1980). A Simple Method for Estimating Evolutionary Rates of Base Substitutions through Comparative Studies of Nucleotide Sequences. *J. Mol. Evol.***16**: 111-120.

K. Korostelev A., Trakhanov S., Laurberg M., Noller H.F., (2006).Crystal Structure of a 70S Ribosome- t RNA Complex Reveals Functional Interactions and Rearrangements. *Cell*. **126**: 1065-1077.

L. Lehmann D., Hettwer H. and Tara Schewski H., (2000). RAPD-PCR Investigations of Systematics Relationships Among four Species of Eels (Teleostei, Anguillidae), Particularly Anguillae and A. rostrata. *J. Mar.Bio.***137**: 195-204.

M. Mau B, Newton M. A., (1997). Phylogenetic Inference for Binary Data on Dendograms using Markov chain. *Journal of computational and graphical statistics.***6**: 122-131.

N. Nagahama T., Sato H., Shimizu M., and Sugiyama J., (1995). Phylogenetic Divergence of the Entomophthoralean Fungi: Evidence from Nuclear 18S Ribosomal RNA Gene Sequences. *Micologia*. **87**: 203-209.

O. Nazar R.N., Sitz T.O., Busch H. (1976). Sequence Homologies in Mammalian 5.8S Ribosomal RNA.

P. Olsen G.J. and Woese C.R., (1993). Ribosomal RNA: A Key to Phylogeny. *FASEB journal.* **7:** 113-123.

Q. Richard G. F., Kerrest A., Dujon B., (2008). Comparative Genomics and Molecular Dynamics of DNA Repeat in Eukaryotes.

R. Saiki K.K., Gelfand D.H., Stoffel S., Scharf S.J., Higuchi R., Horn G.T., Mullis K.B., Erlich H.A., (1989). Primer-Directed Enzymatic Amplification of DNA with a Thermo Stable DNA polymerase.**140**: 7-16.

S. Saiki K.K., Erlich. H. A., (1992). The Design and Optimisation of the PCR: PCR technology-principles and applications for DNA amplification. W. S. Freeman and company, New York. **146**: 7-22.

T. Sambrook and Russell, (2009). Cold Spring Harbor Laboratory Press, New York. Molecular Cloning Vol. **1-2**.

U. Sandy Primrose, Richard Twyman and Bob Old, (2003). Principles of Genetic Engineering, Blackwell Publishing Company.

V. Schierwater B., Streit B., and Wanger G. P., (1994). Molecular Ecology and Evolution; Approaches and Applications. Birkhauser Verlag. Switzerland.

W. Slutmann H., Mayer W. E., Figueroa F., Tichy H., *et al*., (1995). Phylogenetic analysis of cichlid Fishes using Nuclear DNA. *Mol.Bio.* **12**: 1033-1047.

X. Steve Gagnon, David Bourbeau, Roger C. Levesque, (1996). Secondary structures and features of the18S, 5.8S and 26S ribosomal RNAs from the Apicomplexan parasite Toxoplasma gondii. *J. Gene.* **173**: 129-135.

Y. Sullivan J., Joyce P., (2005). Model Selection in Phylogenetics. *Annual Review of Ecology Evolution and Systematics.* **35**: 445-466.

Z. Taggart J.B., Hynes R. A., Prodohl P. A. and Ferguson A., (1992). An Amplified Protocol for Routine Total DNA Isolation from Salmonid Fishes. *J. Fish Bio*. **40**: 963-965.

AA. Tamura K, Dudley J., Nei M. and Kumar S., (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0, *Mol. Bio. Evol.* **8**: 1596-1599,

BB. Teresa K. Attwood., (2010). Essentials of Bioinformatics. Pearson.

CC. Wildman D. E., (2009). Molecular Phylogenetics and Evolution, *Elsevier*. **53**: 694-702

DD. Wilson K. and Walker J., (2010). Practical Biochemistry, Principle and Techniques, Cambridge University Press.

- Yi S., Yi- Cheng Z., Wen- Qin S., Run Sheng Z. and Rui- Yang C., (1999). Detection of Genetic Relationship among Four Artemia Species using Randomly Amplified Polymorphic DNA (RAPD). *International Journal of Salt Lake Research.* **8**: 139-147.