CHAPTER 2:

REVIEW OF LITERATURE

Fishes exhibit greater species diversity than any other group of vertebrates with over 32,000 recorded species (Froese and Pauly, 2012). Fishes do not represent a monophyletic group and therefore the evolution of fish is not studied as a single event. Early fish from the fossil records are represented by a group of small, jawless, armoured fish known as Ostracoderms. Over 10000 fish species live in fresh water (Smith *et al.*, 2000); approximately 40% of global fish diversity and one quarter of global vertebrate diversity. The threats to global freshwater biodiversity can be grouped under five interacting categories: overexploitation; water pollution; flow modification; destruction or degradation of habitat ; and invasion by exotic species (Allan and Flecker, 1993; Jackson Carpenter and Dahm, 2001; Malimqvist and Rundle, 2002; Naiman and Magnuson, 1995; Naiman and Turner, 2000; Postel and Ricther, 2003; Rahel, 2002; Revenga Campbell and Abell, 2005).

The exploration of fish diversity is very problematic because of a great range of divergence. Taxonomic tools and expertise is lacking in this regard. The taxonomists working in developing countries like India which harbours a lion's share (>95%) of described fish species were facing major problems while identification of these fishes like difficulty in accessing basic taxonomic information such as species descriptions. And when taxonomic keys and descriptions are available, they were very rarely revised and so they were often inadequate to identify specimens to the species level. This problem in exploring fishes has been imposing great threat to global fish biodiversity.

FishBase is a global fish database developed by International Centre for Living Aquatic Resource Management (ICLARM) and Food and Agricultural Organization (FAO). It provides information on nomenclature, distribution, ecology, reproduction growth and mortality. FishBase is the largest and most extensively accessed online database on adult finfish on the web. It has evolved into a dynamic and versatile ecological tool which is widely cited in scholarly publications (Stergiou and Tsikliras, 2006). FishBase is now become the largest and most accessed online database for fish in the world (Palomares and Bailly, 2011).

2.1 Ornamental fish inventory in India and North east India perspective

Extensive studies on freshwater fishes regarding taxonomy, biology and aquaculture were carried out by certain researchers namely, Hamilton (1822), Shaw and Shebbeare (1937), Hora (1951 and 1953), Misra (1959), Menon (1974 and 1999), Dey (1973), Jayaram (1981 and 1999), Sen (1982 and 1985), Talwar and Jhingran (1991), Nath and Dey (1977 and 2000), and Dey and Kar (1989). The marine fishes were distributed along the coastal sites and so far a total of 1740 species of marine fishes were described from India. Besides the marine fishes there were a total of 862 species of freshwater fish species described so far of which 672 species are of primary freshwater group (Froese and Pauly 2013). Around 33% of Indian freshwater fishes were distributed along the North-Eastern Region.

In 1982, Ghosh and Lipton reported 172 species of fishes with reference to their economic importance while Sen (1985), reported 187 species of fishes from Assam and its environs. Singh in 1994 compiled a list of 230 species of fishes from north-eastern region of India. Nevertheless, Nath and and Dey (1997), recorded 131 species of fishes from the drainages in Arunachal Pradesh alone. Further, according to Sen 1982, 1985 and 2000, of the 806 species of fishes inhabiting the freshwaters of India (Talwar and Jhingran 1991), North-Eastern region alone was found to represent 33.13% of the total Indian freshwater fish species. Regarding ichthyofauna of Mizoram, Tripura and Barak drainage in Assam Kar and Sen (2007) has given a good output.

In India, about more than 288 exotic species of ornamental fish have been recorded and more than 200 species of these are freshwater fishes and are bred in different parts of India and still other species are imported as fry (Mahapatra *et al.*, 1999). A total of 400 species (250 indigenous fresh water fish and 150 marine fish) have been identified as potential and suitable for tropical climate of India (Rana, 2007; Sane, 2007).

According to Ramachandran (2002), the riverine ornamental fisheries of the Western Ghats in the South Indian state of Kerala is widely considered to be one of the best in the

world with the presence of more than 100 species. Kurup B (2006) reported 106 species of ornamental fishes solely from the state of Kerala.

In a pioneering study on ornamental fishes done by Central Marine fisheries Research Institute, Kochi, India 165 marine ornamental fish species were identified and reported from Lakshadweep and documented in the publication "Marine ornamental fish resources of Lakshadweep". In Gulf of Mannar 113 marine ornamental finfish species were recorded and their biomass was assessed. Andaman and Nicobar islands contributes more than 250 ornamental fish species out of total 1200 species identified (Pathak *et al*., 2012).

After a thorough study made by Ponniah and Sarkar (2006), 28.44% of the total fishes found in North East India are considered as ornamental. A total of 58 potential ornamental fishes have been reported under 42 different genera and 21 families in the Bordoibam Bilmukh Bird Sanctuary, North Eastern India (Dutta NN *et al.,* 2012). Unfortunately no research has been done on the inventory of ornamental fishes from the state of Manipur although many works on endemic fishes has been done by Vishwanath *et al*., (1985, 1998, 2000, 2004, 2005, 2007, 2012). They were exploring the endemic fishes of Manipur not taking into account their ornamental values but fortunately they are regarded as highly prized ornamental fishes in the global market (MPEDA, 2011). Recent works on the inventory of these endemic fishes include works on Fishes of the genus *Glyptothorax* Blyth from Manipur, India with description of three new species (Linthoingambi and Vishwanath, 2007), Fishes of the genus *Osteobrama heckel* of North Eastern India (Vishwanath and Shantakumar, 2007), *Barilius profundus*, a new cyprinid fish (Teleostei: Cyprinidae) from the Koladyne basin, India (Dishma and Vishwanath, 2012), *Mystus ngasep*, a new catfish species (Teleostei: Bagridae) from the headwaters of Chindwin drainage in Manipur, India (Darshan *et al*., 2011) and so on.

2.2. Progress of molecular taxonomy in ichthyodiversity

Species identification by molecular analysis has been utilized for many years. Initially, allozyme differences were used (Avise, 1975), followed by mtDNA

examination (Avise, 1994). Genomic approaches to taxon diagnosis exploit diversity among DNA sequences to identify organisms (Kurtzman, 1994; Wilson, 1995), a bacterial identification system using srRNA sequences have been developed proved to be effective (Busse *et al*., 1996). DNA based methods have several advantages over their protein-based counterparts because DNA is less sensitive to degradation (Hanner, 2005) and can be accessed in all stages from egg to adult. Furthermore, synonymous mutations can be recognized in sequencing approaches, and polymerase chain reaction (PCR) amplification protocols make it possible to analyse minute amounts of tissue. Perhaps most importantly, DNA sequence data are easier to replicate and interpret across laboratories. It is not a coincidence that DNA barcoding has developed in concert with genomics-based investigations. DNA barcoding and genomics share an emphasis on large scale genetic data acquisition that offers new answers to questions previously beyond the reach of traditional disciplines. DNA barcode consist of a standardized short sequence of DNA that in principle should be easily generated and characterised for all species on the planet.

In 2003, Hebert *et al*., from the University of Guelph, Ontario, Canada, first established that the mitochondrial gene *COI* can serve as the core of a global bio identification system for animals by discriminating 200 closely allied species of Lepidopterans and proposed the compilation of public library of DNA barcodes that would be linked to named specimens. Hebert *et al*., in 2004 also sequenced DNA barcodes of 260 of the 667 bird species that breed in North America. They found that every single one of the 260 species had a different *COI* sequence. The sequences were either identical or were most similar to sequences of the same species.

DNA barcoding has several advantages over previous methods. One advantage is its availability. The standard DNA barcode region, a fragment of *COI*, is very efficient for species identification and has good discrimination power for most animal groups. The universal primer, originally designed for marine invertebrates, proved effective for many animal phyla (Folmer *et al*., 1994; Hebert *et al*., 2003; Busenlehner *et al*., 2004; Hebert *et al.,* 2004).

Thereafter, the Barcode of Life project was proposed to promote DNA barcoding as a global standard for sequence-based identification of eukaryotes. In 2004, this project was formally initiated by the establishment of the Consortium for the Barcode of Life (CBOL), which aims to develop a standard protocol for DNA barcoding and to construct a comprehensive DNA barcode library.

The next major study into the efficacy of DNA barcoding was focused on the Neotropical skipper butterfly, *Astraptes fulgerator* at the Area Conservacionde Guanacaste (ACG) in North-Western Costa Rica. This species was already known as a cryptic species complex. Hebert *et al*., in 2004 sequenced the *COI* gene of 484 specimens from the ACG. The studies finally come with conclusion that *Astraptes fulgerator* consists of 10 different species in North-Western Costa Rica.

In 2005, Lambert *et al.,* examined the possibility of using DNA barcoding to assess the past diversity of the earth's biota. The *COI* gene of a group of extinct ratite birds the moa, were sequenced using 26 sub-fossil moa bones. As with Hebert's results, each species sequenced had a unique barcode and intraspecific *COI* sequence variance ranged from 0 to 1.24.

Marine biologists have also considered the value of the technique in identifying cryptic and polymorphic species and have suggested that the technique may be helpful when associations with voucher specimens are maintained (Saunders, 2005).

In 2006, Smith *et al*., examined whether a *COI* DNA barcode could function as a tool for identification and discovery for the 20 morphospecies of *Belvosia* parasitoid flies (Tachinidae). Barcoding not only discriminated among all 17 highly host-specific morphospecies of *Belvosia*, but it also suggested that the species count could be as high as 32 by indicating that each of the three generalist species might actually be arrays of highly host-specific cryptic species. However, in 2007, Whitworth *et al*., reported that flies in the related family Calliphoridae could not be discriminated by barcoding. They investigated the performance of barcoding in the fly genus *Protocalliphora*. Assignment of unknown individuals to species was impossible for 60% of the species, and if the technique had been applied, as in the previous study, to identify new species, it would

have underestimated the species number in the genus by 75%. They attributed the failure of barcoding to the non-monophyly of many of the species at the mitochondrial level.

In 2007, Hajibabaei *et al*., compared the goals and methods of DNA barcoding with those of molecular phylogenetics and population genetics, and suggest that DNA barcoding can complement current research in these areas by providing background information that will be helpful in the selection of taxa for further analyses.

In 2007, Min and Hickey showed that short sequences of *COI* DNA barcode can yield important, and surprisingly accurate, information about the composition of the entire genome. Thus, for un-sequenced genomes, the DNA barcodes can provide a quick preview of the whole genome composition. Borisenko *et al*., 2009 however, proposed efficient logistics of pre-laboratory specimen processing and seamless interfacing with molecular protocols for building a global library of DNA barcodes.

By then, the Barcode of Life project entered a new phase with the launch of the International Barcode of Life project (iBOL; International Barcode of Life 2010a). The iBOL is a huge international collaboration of 26 countries that aims to establish an automated identification system based on a DNA barcode library of all eukaryotes.

The first description of a new species using a DNA barcode from the holotype was by Brown *et al.*, 2003, who used this method to describe a new species of Xenothictis (Lepidoptera:Tortricidae). Since then, many new species have been described with DNA barcodes from the holotype or paratypes, not only in arthropods, but also in other animals (Burns *et al*., 2007; Badek *et al*., 2008; Dabert, 2008; Yassin *et al*., 2008; Yoshitake *et al*., 2008; Adamski *et al*., 2009).

In 2010, Costa and Carvalho explained the prospects of barcode of life initiative (BOLI) where they mentioned- "while genome projects yield spectacular insights into molecular evolution, they have targeted only a few species. In contrast, the Barcode of Life Initiative (BOLI) proposed a horizontal approach to genomics, examining short, standardized genome segments across the sweep of eukaryotic life, all 10 million species".

`In 2011, Meiklejohn *et al*., brought forward a wonderful practical application of DNA barcoding by sequencing of a 658-bp 'barcode' fragment of *COI* gene of forensically important Sarcophagidae (Diptera) from 85 specimens, representing 16 Australian species from varying populations. All species were resolved as reciprocally monophyletic, except Sarcophaga dux. The *COI* 'barcode' sequence was found to be suitable for the molecular identification of the studied Australian Sarcophagidae.

In 2011, McGowin *et al*., focussed the importance of *COI* DNA barcode in identification of two marine turtle leeches (*Ozobranchus margoi* and *Ozobranchus branchiatus*. Using morphological taxonomy combined with distance- and character-based genetic sequence analyses, this study has established a DNA barcode for both species of *Ozobranchus* spp. leech and has shown that it can be applied successfully to the identification of leeches at earlier stages of development when morphological taxonomy cannot be employed.

Again in 2011, McFadden *et al*., however, found limitations of DNA barcoding for identification of an anthozoan cnidaria (Octocorallia) where they explained- "although far from perfect for species identification, a *COI* + igr1 + msh1 barcode nonetheless represents a valuable addition to the depauperate set of characters available for octocoral taxonomy".

In 2011, Bell *et al*., also focussed importance of DNA barcoding in European liverwort species from the genus *Herbertus*. The study was undertaken using three plastid (*matK*, *rbcL* and *trnH-psbA*), and one nuclear (*ITS*) marker. The DNA barcode data were effective in discriminating among the sampled species of *Herbertus* and contributed towards the detection of a previously overlooked European *Herbertus* species, described here as *H. norenus* sp. nov. The species showed clear-cut differences in DNA sequence for multiple barcode regions and is also morphologically distinct. The DNA barcode data were also useful in clarifying taxonomic relationships of the European species with some species from Asia and North America. In terms of the discriminatory power of the different barcode markers, ITS, was the most informative region, followed closely by *matK*. All species were distinguishable by ITS alone, *rbcL*+ *matK* and various other multi marker combinations.

Bennett *et al.,* in 2011 carried out DNA barcoding of an invasive mammal species, the small Indian mongoose (*Herpestes javanicus*; E. Geoffroy Saint-Hillaire, 1818) in the Caribbean and Hawaiian Islands. The work demonstrates the utility of using DNA barcoding approaches with mtDNA cytochrome b to discriminate between the two species and other sympatric members of the genus *Herpestes* (*Herpestesnaso*, *Herpestes urva*, and *Herpestes edwardsii*). Using the diagnostic DNA positions, specimens of non-native populations of the small Indian mongoose from the Caribbean and Hawaiian Islands was assigned to their species of origin. A single diagnostic site accomplishes the identification of *H. javanicus* versus *H. auropunctatus*. Then results came with the conclusion that the non-native mongoose populations from the Caribbean and Hawaiian Islands are *H. auropunctatus*, and not *H. javanicus.* Bitanyi *et al.,* in 2011 evaluated the effectiveness of *COI* barcode in species identification of Tanzanian antelope species. A 470 base-pair region of the *COI* gene was examined in 95 specimens representing 20 species of antelopes, buffalo and domestic Bovidae. All the Tanzanian species showed unique clades. Further they demonstrated that even short *COI* fragments can efficiently identify antelope species.

Bucklin *et al*., in 2011 reviewed the role of DNA barcoding in the study of marine metazoa, with concern that more than 230,000 known species representing 31 metazoan phyla populate the world's oceans and perhaps another 1,000,000 or more species remain to be discovered.

In 2012, Lauchli *et al*., developed a DNA barcoding workflow capable of processing potentially large sponge collections and is routinely used for the Sponge Barcoding Project with success. Sponge specific problems such as the frequent co-amplification of non-target organisms have been detected and potential solutions are currently under development. The initial success of this innovative project have already demonstrated considerable refinement of sponge systematics, evaluating morphometric character importance, geographic phenotypic variability, and the utility of the standard barcoding fragment for Porifera (despite its conserved evolution within this basal metazoan phylum). The campaign of barcoding still continues in with the involvement of more and more organisms (Bell *et al*., 2011; Clare *et al*., 2011; Janzen *et al*., 2011) to reveal the actual level of diversity among them.

Some researchers have envisioned "DNA taxonomy", a concept of adopting DNA sequencing as a central criterion for taxonomic decisions and descriptions, and have proposed using DNA barcodes as the standard method of analysis (Blaxter, 2003; Tautz *et al*., 2003; Vogler *et al*., 2007). However, there is concern over adopting one specific sequence region as the only criterion for taxonomic studies (Lipscomb *et al*., 2003; DeSalle *et al.,* 2005; Rubinoff, 2006). In addition, it is quite apparent that the DNA barcode itself is not a new species concept (i.e. a species cannot be defined based on the barcode only); neither does it provide enough information to describe unknown specimens as a new species. The results of barcoding can only suggest new species candidates (Brown *et al*., 2003; Hajibabaei *et al.,* 2006; Hajibabaei *et al*., 2007; Waugh, 2007) as well as other valuable supporting information (e.g. distribution, life history, host plants) for taxonomic studies e.g. integrative taxonomy (Dayrat, 2005; Yoshitake *et al.*, 2008; Schlick-Steiner *et al.*, 2010). Species descriptions using barcodes based on type specimens will become more common and important in the near future.

2.3 Application of DNA barcoding for decoding of Ornamental fish

In 1991, Bartlett *et al*., were among the first to use mtDNA sequencing for fish identification, showing that *cytochrome b* sequences could discriminate four species of tuna (*Thunnus* spp.). They subsequently proposed forensically important nucleotide sequences (Bartlett *et al.*, 1992) as a means of identifying fishes.

A large community of scientists joined forces in 2005 to launch the Fish Barcode of Life (FISH-BOL) campaign (Lundberg *et al*., 2000) to meet these needs of an accurate inventory of species and a more scalable and cost effective approach to their reliable identification at any life-history stage.

In 2005, Ward *et al*., had sequenced (barcoded) for a 655 bp region of the mitochondrial *COI* of 207 species of fish. All species could be differentiated by their *COI* sequence, although single individuals of each of two species had haplotypes characteristic of a congener.

Dawnay *et al*., in 2007 introduced the application of DNA barcodes to forensics for wildlife crime investigation which previously routinely involved genetic species

identification based on DNA sequence similarity. However, this work was hindered by a lack of authenticated reference DNA sequence data resulting in weak matches between evidence and reference samples. The introduction of DNA barcoding has highlighted the expanding use of the mtDNA gene, *cytochrome c oxidase I* (*COI*), as a genetic marker for species identification. They assessed the COI gene for use in forensic analysis following published human validation guidelines. Validation experiments investigated reproducibility, heteroplasmy, mixed DNA, DNA template concentration, chemical treatments, substrate variation, environmental conditions and thermo cycling parameters. Sequence similarity searches using both GenBank BLASTn and BOLD search engines indicated that the *COI* gene consistently identifies species where authenticated reference sequence data exists. Where misidentification occurred the cause was attributable to either erroneous reference sequences from published data, or lack of primer specificity. Although amplification failure was observed under certain sample treatments, there was no evidence of environmentally induced sequence mutation in those sequences that were generated. A simulated case study compared the performance of *COI* and cytochrome b mtDNA genes.

The status of *E. eggvinii*, was reassessed in 2007 using 21 meristic and 32 morphometric characters were analysed for a total of 83 specimens of *E. spinosus* and *E. eggvinii*. Mitochondrial (*COI, COII and cyt-b*) and nuclear (Tmo-4C4) genes were also sequenced for both species, along with *Eumicrotremus derjugini*. The results indicate that although *E. spinosus* and *E. eggviniiare* clearly separated by a considerable number of morphological characters, they in fact constitute a single, sexually dimorphic species. Thirteen specimens of *E. eggvinii* (including the holotype) and 59 *E. spinosus* could be sexed; all individuals of E. eggvinii turned out to be males and all E. spinosus were females. Identical DNA sequences were found in all *E. eggvinii* and *E. spinosus* for *COI, COII* and *Tmo-4C4*, and a single shared synonymous substitution found in cyt-b. In contrast, *E. spinosus, E. eggvinii* and *E. derjugini* differed by 5·9% for *COI* and *COII*, 1·2% for Tmo-4C4 and 8·3% for *cyt-b*. The genus *Eumicrotremus* comprises 16 lumpsucker species distributed in the Arctic and northern Atlantic and Pacific Oceans. The most common species in the North Atlantic is *Eumicrotremus spinosus*, described in 1776, and characterized partly by numerous bony tubercles on the head and body.

Another Atlantic species, *Eumicrotremus eggvinii*, described in 1956, remained known only from a single specimen until additional specimens were recently recovered (Byrkjedal *et al*., 2007).

The pioneering work on decoding the ornamental fishes was done by Steinke D, Zemlak TS and Hebert PDN in the year 2009. They developed a barcode library, though incomplete, of various ornamental fishes being traded and being imported in North America which already provides a new species identification tool for the ornamental fish industry opening a realm of applications linked to collection practices, regulatory control and conservation (Steinke D, Zemlak TS, Hebert PDN, 2009).

In 2009, Cohen et al., practically applied DNA barcode in identifying commercially available puffer fish. In 2007, two individuals developed symptoms consistent with tetrodotoxin poisoning after ingesting home-cooked puffer fish purchased in Chicago. Both the Chicago retailer and the California supplier denied having sold or imported puffer fish but claimed the product was monkfish. However, genetic analysis and visual inspection determined that the ingested fish and others from the implicated lot retrieved from the supplier belonged to the family Tetraodontidae. Tetrodotoxin was detected at high levels in both remnants of the ingested meal and fish retrieved from the implicated lot. The investigation led to a voluntary recall of monkfish distributed by the supplier in three states and placement of the supplier on the U.S. Food and Drug Administration's Import Alert for species misbranding. This case of tetrodotoxin poisoning highlights the need for continued stringent regulation of puffer fish importation by the U.S. Food and Drug Administration, education of the public regarding the dangers of puffer fish consumption, and raising awareness among medical providers of the diagnosis and management of food borne toxin ingestions and the need for reporting to public health agencies.

In 2010, Botti described a method using existing PCR and sequencing methodologies to detect mitochondrial DNA polymorphisms in complex matrices such as foods. The reported application allowed the discrimination among 17 fish species of the Scombridae family with high commercial interest such as mackerels, bonitos and tunas which are often present in processed seafood. The approach can be easily upgraded with the release

of new genetic diversity information to increase the range of detected species (Botti *et al.*, 2010). In 2011, Asgharian *et al*., with regard to practical applicability of DNA barcoding generated a large scale datasets of mitochondrial *COI* gene of fish of the Nayband National Park in the Persian Gulf.

Khedkar *et al.*, in the year 2011 have undertaken DNA barcoding of ornamental fishes available in India. A total of 135 specimens of ornamental fishes from commercial market were collected. Most of them were apparently traded with different local names and different price. All collected fishes were classified morphologically; around 15 fishes were very difficult for morphological based identification. For all 135 fishes, DNA was obtained from a small fin clip, COI gene amplified and sequences obtained; after blast it was found that many fishes look morphologically different but are the same. A total of 53 species can be validated from the available data with BOLD (Khedkar *et al*., 2011).

In 2011, Amor *et al*., molecular characterized *Hysterothylacium aduncum* (Nematoda: Raphidascaridae) from different fish caught off the Tunisian coast based on nuclear ribosomal DNA sequences instead of *COI* DNA barcode.

In 2011, Carvalho *et al.*, unveiled a high rate of mislabelling in a commercial freshwater catfish from Brazil through DNA barcoding where they reported on the molecular identification results from processed fish products (i.e. fillets) and whole fishes sold in Brazilian markets under the common name surubim (Pseudoplatystoma spp.). They found DNA barcoding revealed the incorrect labelling of around 80% of all samples analyzed, with mislabelling being more pronounced within fillets rather than whole fish.

In 2011, Benziger *et al.*, resolved the taxonomic ambiguity, and discussed the identity as well as systematic position of the Malabar snakehead, C. diplogramma, using morphological and molecular genetics (mitochondrial 16S rRNA and *COI* gene) information, in addition to making an attempt to understand its phylogenetic relationships and evolutionary biogeography. Both morphological and genetic analyses

support *C. diplogramma* as a distinct and valid species endemic to peninsular India and reveal its importance for conservation.

In 2011, Wong *et al*., developed and evaluated DNA barcodes for use in differentiating United States domestic and imported catfish species. They also suggest that as the United States heightens inspection and regulation requirements for sea food products, DNA barcoding will serve as an important tool in efforts to ensure consumer safety and fair international commerce.

In 2011 again, April *et al*., established a barcode reference library for more than 80% of the named freshwater fish species of North America. This study demonstrates that 90% of known species can be delineated using barcodes. Moreover, it reveals numerous genetic discontinuities indicative of independently evolving lineages within described species, which points to the presence of morphologically cryptic diversity. From the 752 species analyzed, their survey flagged 138 named species that represents many as 347 candidate species, which suggests a 28% increase in species diversity. In contrast, several species of parasitic and non-parasitic lampreys lack such discontinuity and may represent alternative life history strategies within single species.

Becker *et al*., in 2011 provided a 5-year progress report on the campaign and includes an updated "Collaborators' Protocol" (Steinke *et al.*, 2011) to facilitate its continued growth and success. The implementation of standards (Dey *et al*., 1989) is attributed to the overarching success of barcoding (Allan *et al*., 1993) and to this end, the new protocol aims to refine and further advance FISH-BOL best practices for the benefit of the user community. Key to this objective is the widespread adoption of specimen imaging and reporting of identification "confidence levels" as discussed in the new protocol, which also reiterates the importance of a shared informatics workbench, the Barcode of Life Data system (Ratnasingham S, 2007).

The utility of FISH-BOL derives from the contributions of many and varied researchers from around the world who are dedicated to expanding the barcode coverage for global fishes. The accumulating data already support applications of DNA barcoding which reveal market substitution (Naiman *et al*., 1995; Naiman *et al.*, 2000; Jackson *et*

al., 2001) and enhancing our understanding of fisheries exploitation (Holmes *et al.,* 2009; Doukakis *et. al*., 2011). Yet the broad realization of benefits is predicated on a sustained effort to complete the construction of reference sequence library. From Africa (Malimqvist *et al.,* 2002; Rahel, 2002) and Europe (Postel *et al*., 2003), Oceania (Revenga *et al*., 2005) and South America (Carvalho *et al.,* 2011; Pereira *et al*., 2011) a large number of researchers have contributed to the FISH-BOL campaign. Out of almost 30,000 fish species estimated in the world, barcodes for more than 10,000 fish species are currently recorded in the BOLD database.

In 2011, Bucklin *et al.,* calculated an average retrieval of 2% new species in larger fish DNA barcoding studies, and they extrapolated this rate to about 600 overlooked or cryptic species to await discovery through similar studies. From the 31,000 species currently listed in the Catalogue of Fishes, about 4000 have been described new during the past 10 years (2000–2009), with 500 added in 2008 and 300 in 2009 (Eschmeyer *et al*., 2012).

Weigt *et al*., in 2012 represented a DNA barcode data release for 3,400 specimens representing 521 species of fishes from 6 areas across the Caribbean and western central Atlantic regions (FAO Region 31). Merged with their prior published data, the combined efforts result in 3,964 specimens representing 572 species of marine fishes and constitute one of the most comprehensive DNA barcoding ‘‘coverage’’ for a region reported to date.

In 2012, Kadarusman *et al.,* assessed the diversity of the Papua rainbow fishes with DNA barcoding. They sequenced the mitochondrial *COI* gene for 350 specimens belonging to 53 nominal species throughout the Indo-Australian archipelago. Unexpected levels of cryptic diversity and endemism were detected since additional cryptic lineages were detected in several watersheds from the Vogelkop and the Lengguru massif. DNA barcoding supports the presence of nearly 30 evolutionary lineages among the 15 nominal species sampled in the Vogelkop and all these lineages are endemic to a single lake or watershed. This result highlights that the diversity of the family has been largely underestimated and urges for the identification of conservation priorities in Papua.

Costa *et al.,* in 2012 proposed a ranking system to attribute a confidence level to species identifications associated with DNA barcode records from reference libraries of DNA barcodes (RLDB). The increasing availability of reference libraries of DNA barcodes (RLDB) offers the opportunity to the screen the level of consistency in DNA barcode data among libraries, in order to detect possible disagreements generated from taxonomic uncertainty or operational shortcomings.

In 2012, Collins *et al*., presented a barcoding approach for ornamental cyprinid fishes by expanding current barcode reference libraries; assessing barcode congruence with morphological identifications under numerous scenarios (e.g. inclusion of GenBank data, presence of singleton species, choice of analytical method); and providing supplementary information to identify difficult species. DNA barcoding offered a potentially attractive tool for quarantine inspection, as scrutinised for aquarium fishes. This was thus an efficient solution to the problem of poor regulation of international trade in ornamental fishes, which poses risks to both biodiversity and economic activity via invasive alien species and exotic pathogens. Border security officials need robust tools to confirm identifications, often requiring hard-to-obtain taxonomic literature and expertise in this context species identification through DNA barcoding provided an effective solution.

The FISH-BOL campaign currently has barcoded for the *cytochrome c oxidase subunit I* (*COI*) gene about 8,000 of the 31,000 fish species currently recognised. This includes the great majority of the world's most important commercial species. Results thus far show that about 98% and 93% of marine and freshwater species, respectively, are barcode distinguishable. One important issue that needs to be more fully addressed in FISH-BOL concerns the initial misidentification of a small number of barcode reference specimens. This is unsurprising considering the large number of fish species, some of which are morphologically very similar and others as yet unrecognised, but constant vigilance and ongoing attention by the FISH-BOL community is required to eliminate such errors. Once the reference library has been established, barcoding enables the identification of unknown fishes at any life history stage or from their fragmentary remains. The many uses of the FISH-BOL barcode library include detecting consumer fraud, aiding fisheries management, improving analyses including food web syntheses, and assisting with taxonomic revisions (Ward, 2012).