We propose to develop a phylogeny for the Cypriniformes, the largest clade of entirely freshwater fishes with 3,285 described, and as many as 2,600 undescribed species. This group contains the minnows, suckers, loaches, river loaches and algae eaters. Many of these fishes have economic and scientific importance. Several species of large carp are an important protein source in Asia, and paradoxically, are nuisance, invasive species in North America. The zebrafish is a model organism for developmental biologists; the fathead minnow is used as an indicator species in environmental work; and goldfish are widely studied by physiologists. Many minnows and loaches are kept by aquarists, a hobby that has vast educational and economic impacts. Our project is international in scope with 27 collaborators, including some of the most prominent cypriniform systematists from across the globe. The group includes paleontologists, neontologists, developmental biologists, and molecular systematists, whose collective expertise encompasses the range of diversity within Cypriniformes.

Intellectual Merit: We will develop a phylogenetic hypothesis based on these diverse data for 1000 species of this large and important clade of fishes. We will collect genomic data for a set of nuclear and mitochondrial genes, including a subset of complete mitochondrial genomes as well as morphological data from fossil and extant species. The morphological data will include developmental characters from 30 exemplar species. The phylogeny, including the developmental information will broaden the scope of work based on the zebrafish and other models. We will test hypotheses of molecular and morphological evolution. Inclusion of fossil and molecular data will allow us to calibrate timing of events in the evolution of cypriniform fishes including morphological and genomic diversification (such as polyploidy), as well as vicariant and dispersal events associated with biogographic history of this group. We will describe new species and generate information on distribution expanding our understanding of biodiversity. Major research questions include the following 1) Are the major cypriniform clades monophyletic? 2) How do characters of phylogenetic significance develop from protein-gene interactions to fully formed adult morphology. 3) Are there constraints imposed on morphology that generate convergence which inaccurately reflect underlying relationships? 4) How does polyploidy impact morphological diversity? 5) What is the age of Cypriniformes and major groups within this clade. 6) What is the relationship between the phylogenetic history of this ancient lineage and known Earth history? Several other important objectives related to these and expanding beyond these are outlined.

Broader Impacts: Our study has significant broader impacts for the general public and other researchers. We will train graduate and undergraduate students in phylogenetic systematics, development, and morphology. We will support international collaboration and research. We will broaden the scope and impact of research on model species such as the zebrafish, fathead minnow, and goldfish. We are collaborating with the NSF funded All Catfish Species Inventory to describe and inventory diversity of freshwater fishes. We will develop a Web Portal that will communicate our work, and provide information on biology, distribution, and economic impact of these fishes to the general public. This portal will be linked with other databases including FishBase and the Tree of Life web page, providing a wealth of accessible information on these fishes. We plan collaborations with natural history museums and aquariums to develop exhibits, and an educational program for K-12 students. Finally, we are working with professionals on a PBS-oriented television documentary about the discovery of diversity, the AToL initiative and how “we” reconstruct such a Tree of Life, the fascinating cultural aspects surrounding cypriniform fishes, their impact as invasive species, and how modern-day science in a global environment must involve interdisciplinary and highly connected networks of broadly trained scientists.
II. Introduction

Freshwater ecosystems contain only about 0.01% of the Earth’s water but at least 45% of all fish species. Freshwater systems are arguably among the most endangered ecosystems on Earth; growing human populations have had a negative impact on freshwater biodiversity and political conflicts over water (18, 85) will only increase the problem. The need to document the biodiversity of freshwaters is critical if we are to evaluate the risk factors for species conservation and ecosystem restoration. Cypriniformes is the most diverse groups of freshwater fishes on the planet, with an estimated 3,285 described species contained in 280 genera and five families (www.fishbase.org; we estimate that over 2,600 remain undescribed). We propose an international, collaborative project to discover and describe diversity, resolve basal and apical relationships within the Cypriniformes based on varied data sets, and produce a classification of Cypriniformes, earth’s most diverse clade of freshwater fishes.

Cypriniformes contains the minnows, suckers, loaches, river loaches, and algae eaters. These fishes are found in a vast array of environments, from subterranean aquifers to streams in the Himalayas; from brackish estuaries to inland lakes; and from icy mountain streams to thermal springs. This broad range of environments is mirrored by the extensive morphological diversity. Species delineation sometimes poses a problem and there is certainly a huge amount of undiscovered diversity within this taxon (ca. 2,600 undescribed species). Cypriniforms are absent from Australia and South America but widespread through Africa, Asia, Europe and North America; these fishes usually comprise most of the diversity in aquatic environments. These fishes are very important economically and most people are familiar with some species or aspects of their biology. Indeed, these fishes are common in the pet trade; there are few freshwater community tanks that do not contain brightly colored loaches and minnows, valued for their appearance and ease of maintenance. The first pet many of us owned was a goldfish (Carassius auratus), a cyprinid that has been cultured for its beauty for thousands of years. Fly fishermen travel from around the world to Asia to try their luck landing the mahseer (Tor putitora), a 2.7 meter-long cyprinid that belies the name minnow. In Asia, cypriniforms are an important source of protein and have been maintained in complex polyculture systems for over two-thousand years. Several species have been introduced outside their range and are now considered invasive nuisance species, the best known of these is the common carp (Cyprinus carpio). Considerable resources are devoted to understanding the biology of such invasive species with the hope of controlling their populations and minimizing their impacts on ecosystems. Cypriniform fishes are also increasingly important as research subjects. Goldfish have been widely studied by physiologists, particularly with regard to hormonal and pheromonal impacts on behavior. The fathead minnow (Pimephales promelas) is widely used for environmental testing and is increasingly important in studies of behavior. The zebra danio (Danio rerio) is a model species valued for its use in studies of evolution and development.

Many freshwater aquatic species are precipitously declining around the world, particularly in developing Third World nations due to habitat loss, degradation, and impoundments brought about by rapid human population growth (95, 96). Given the importance of these fishes to various aspects of human society, especially aquaculture and the aquarium trade, the time has come for an international, collaborative effort to document the total amount of diversity within the group, how this diversity is spread across the surface of the planet and how component species are related to one another.

WHY A COLLABORATIVE PROJECT?

"Projects for Assembling the Tree of Life are expected to be ambitious, large scale, and to involve multiple investigators from multiple disciplines, likely from multiple organizations, and to include training, outreach, and dissemination components” (NSF Program Solicitation 04-526). This project extends from ongoing collaborations among several of the PIs. The first stage produced molecular phylogenetic hypotheses covering most major lineages of North American Cyprinids (89-93). This work was originally funded by NSF "Phylogenetic Relationships of North American Cyprinid genera as evidenced by morphology, behavior, and mtDNA sequences” (DEB-9307132). Three additional proposals, “Collaborative Research: Evolution of Phoxinid Trophic Morphology” (AMS, RLM), “Phylogeny and classification of the subfamily Catostominae (Cypriniformes: Catostomidae) based on
nuclear and mitochondrial DNA sequences and morphological characters” (PMH), and “Systematics of Fishes of Subfamily Ictiobinae (Teleostomi: Catostomidae)” (HLB), are either active (HLB) or in review. These studies and other associations among the PIs and with other collaborators have lead to numerous publications, presentations, and educational opportunities that have advanced our knowledge of these fishes; further advancement in our understanding of this diverse clade must be at a larger scale with multiple labs and PIs, and at an international scope. We plan to follow the current proposal with another large scale, collaborative Planetary Biotic Inventory proposal to document all diversity within the Cypriniformes. The group is large, widespread, and external morphology is deceptive, causing difficulties with identification and classification. There have been few attempts to conduct large-scale overviews of cypriniform systematics. Most studies with large taxonomic scope, included very restricted numbers of terminal taxa (87) or are restricted to individual families (13, 15, 83). The AToL initiative provides an opportunity to bring together a pool of researchers with extensive expertise in cypriniform diversity and systematics. Researchers will collaborate on character selection, description, and analysis. As well, we will collaborate in describing undiscovered diversity within the cypriniform fishes.

We have gathered some of the most prominent cypriniform systematists representing a diversity of analytical approaches. This group includes paleontologists, neontologists, developmental biologists, and molecular systematists, whose collective taxonomic expertise encompasses the range of diversity within Cypriniformes. The PIs include: Arratia, morphological systematics, development, fossil and extant actinopterygians; Aspinwall, African biodiversity, aquaculture of Cypriniformes; Bart, molecular systematics, Catostomidae; Coburn, morphological systematics, development, Cypriniformes; Harris, molecular systematics, Catostomidae, Cyprinidae; Mabee, morphological systematics, development, Cyprinidae, Mabee is also a member of ZFIN project, the developmental zebrafish website (http://www.zfin.org); Mayden, morphological and molecular systematics, Catostomidae, Cyprinidae; Simons, molecular and morphological systematics, Cyprinidae; Wood, molecular systematics, Cyprinidae.

We have also assembled a team of associates and international collaborators: Bogutskaya, morphological and molecular systematics, Cyprinidae; Bohlen, morphological and molecular systematics, cytogenetics, Cobitidae; Chang, morphological systematics, fossil and recent Cyprinidae; Clements, molecular evolution, Catostomidae, Cyprinidae; Doadrio, molecular systematics, Cyprinidae; Fang, Morphological systematics, Cyprinidae; George, web development and management; Golubtsov, morphological and molecular systematics, Cyprinidae; Huanzhang, molecular systematics, Cypriniformes; Kottelat, Taxonomy and distribution of Balitoridae, Cobitidae, Cyprinidae, Gyrinocheilidae; Chang, morphological systematics, fossil and recent Cyprinidae; Miya, molecular systematics, mitochondrial genomes, Cypriniformes; Neely, web development and outreach; Rios, web development and outreach; Saitoh, molecular systematics, mitochondrial genomes, Cypriniformes; Salnikov, morphology, Cyprinidae; He, molecular systematics, Cypriniformes; Zardoya, molecular systematics, mitochondrial genomes, Cyprinidae.

This collaborative effort extends beyond Cypriniformes to the All Catfish Species Inventory – wherein participants of both of our groups will share desperately needed samples (NSF funded; see Page letter).

Our collaborators feel that it is critical that we all work as a group to forward the progress of this effort. As such we must share information, coauthor papers, assist one another in sampling of materials, etc. As a group, we have developed a Cypriniformes Team Charter that we have all agreed upon.

### Cypriniformes Team Charter

All participants in the Cypriniformes AToL project are invited to share authorship on any and all projects detailing the results of this study. Given the magnitude of the data sets we are assembling and the diversity of character sets that can be utilized to address systematic issues within this extremely diverse set of taxa, we anticipate many publications as a direct result of our collective efforts. While we are all interested in cypriniform systematics we will not necessarily all be authors on every manuscript resulting...
from this project. As a general premise, we agree to abide by the ethical principles generally applied to
authorship and publication in any scientific endeavor. In practice, we agree to notify other participants
about the topics of proposed manuscripts through a frequently updated “Proposed Manuscript” page
maintained on the portal. Prior to requesting that a group working on a proposed study add one of us as
an additional author we agree to carefully reflect on the following issues: the topic of the proposed
manuscript, our interest in that particular topic, our contribution to the proposed manuscript with regard to
the successful acquisition of the data or taxa required for the study and whether we have truly made or
can make a substantive contribution to the science involved.

History and Current Status of Knowledge

**Cypriniformes.** – The phylogenetic relationships of Cypriniformes have perplexed scientists at least since
Artedi erected one of the first scientific classifications of fishes, and the systematic and taxonomic “state
of affairs” within the order can be described as ranging from being “fairly well understood” to being
unquestionably “chaotic.” Cypriniformes are a species rich and exceptionally diverse group ranking
second or third among orders of vertebrates. FishBase (www.fishbase.org) lists 3,285 species of
cypriniforms as valid. Among freshwater fishes, only the catfishes (Order Siluriformes) with ca. 2,700
valid species are about as species rich as the Cypriniformes. Unlike Siluriformes, which is divided into 34
families and 436 genera (clade.acnatsi.org/allcatfish/families.html), Cypriniformes is divided into five
families (Cyprinidae, minnows; Catostomidae, suckers; Gyrinocheilidae, algae eaters; Cobitidae, loaches;
Balitoridae (Homalopteridae), river loaches) with 280 genera, with ca. 80% of all species in the
Cyprinidae (70). This skewed classification is based, in part, on historic analyses relying either on
external (e.g., barbel presence/absence and morphology) or pharyngeal teeth characters (47). Although
numerous classifications have been proposed, there have been few rigorous phylogenetic analyses
assessing the monophyly of the various taxonomic rankings within this order. A brief history of
cypriniform classification is given below; emphasis is given to those studies employing cladistic
methodology and that assessed the monophyly of various taxonomic groups.

Hensel (33) presented a comprehensive review of the history of cypriniform classification. In
general, most early works attempted to classify regional faunas and collectively grouped the “carps”
(minnows, loaches, suckers) into either the family Cyprinidae, or later, into the suborder Cyprinioidei.
While all of these works recognized a relationship among the “carps,” they differed substantially in
recognizing groups at the familial, subfamilial or tribal levels. Bleeker (4, 5) and Gunther (27) provided
early reviews of the Cyprinoidei, recognizing many groupings that are retained in current classifications
as subfamilies. Gill (23, 24) divided the Eventognathi (carps) into five families: Catostomidae,
Cyprinidae, Cobitidae, Homalopteridae, and Knerridae (Order Gonorrhynchiformes). He subsequently
recognized the algae eaters (Gyrinocheilidae) as an independent family (25). Subsequent workers (e.g.,
Hora, Nichols, Mori) used the familial classification of Gill, minus Knerridae, but variously elevated or
subsumed ‘lineages’ within these families to subfamilial or tribal rankings. Greenwood et al. (26)
proposed the classification of Cypriniformes most commonly used today (although the classification of
*Psilorhynchus* varies as either a monotypic family or a subfamily within Cyprinidae on the author).

Wu et al. (105) presented the first cladistic examination of Cypriniformes relationships, although only
Characiformes was used as an outgroup (see http://bio.slu.edu/mayden/cypriniformes/home.html for
previous phylogenetic hypotheses). Although dichotomous relationships were depicted in their
phylogeny, examination of the distribution of synapomorphies indicates that there are no synapomorphies
supporting the Catostomidae plus Gyrinocheilidae clade; two synapomorphies did support, however, an
unresolved relationship among catostomids, gyrinocheilids, and cobitids. Harris & Mayden (29) resolved
the Cobitidae sister to a clade of Gyrinocheilidae plus Catostomidae based on mtDNA 12S & 16S rDNA
sequences, supporting Wu et al.’s depicted hypothesis.

In contrast, both Sawada’s (83) analysis of the Cobitoidea (Cobitidae plus Homalopteridae) (see
above URL) and Siebert’s (86) analysis of cypriniform relationships (see above URL) yielded
Gyrinocheilus as sister to a clade containing Catostomidae plus (Cobitidae and Homalopteridae). Both
Tree of Life: Systematics of Cypriniformes

III. Research Questions and Objectives of the Cypriniformes AToL Project

“Tree of Life projects that are taxon-oriented will focus on phylogenetic resolution of large lineages or clades” (NSF Program Solicitation 04-526). The primary goal of this Tree of Life Project is to discover and describe diversity, resolve basal and apical relationships within the Cypriniformes based on varied data sets, and produce a classification of Cypriniformes. To achieve this goal, this large, collaborative project brings together cypriniform biologists from around the globe. It includes comparative morphologists, including paleontologists and neontologists; developmental biologists; and molecular systematists. In addition, there is a very close association with the All Catfish Species Inventory Project (see Page letter) that includes reciprocal exchanges of field collections, as well as morphological and molecular datasets for comparative studies. We will sample 1000 species contained in 350 genera classified into five families plus outgroups (the number of genera includes subgenera likely to be elevated and outgroup taxa). All data will be checked on actual specimens, including morphological characters described in the cypriniform literature. We will develop a dynamic Web Portal that will contain digital images of all morphological characters and their codings for all included taxa. This portal will also contain aligned molecular data that will be linked to photographs of cataloged voucher specimens. The portal provides a totally interactive environment for communication among PIs regarding diversity and classification, characters, topical “threaded chats”, and a location where ALL data will be uploaded/downloaded by participants around the world. We have included developmental biologists who

Smith (94) provided the first, all-encompassing phylogenetic analysis of catostomid relationships. In his methods, Smith discussed numerous morphological analyses supporting various cypriniform lineages as potential sister taxa to catostomids. Despite examining numerous cypriniform taxa, however, only Leptobotia (Cobitidae) and Cyprinus (Cyprinidae) were included as outgroups in his data matrix. Smith’s phylogeny depicts a trichotomy among Leptobotia, Cyprinus, and Catostomidae, undoubtedly due to limited taxon sampling among potential outgroup taxa.

Clements et al. (12) provide the most recent hypothesis of cypriniform relationships. Based on Growth Hormone sequences, their analysis resolved a well-supported clade of Cobitidae sister to Catostomidae plus Cyprinidae; this latter clade was also proposed by Uyeno and Smith (100) based on karyotype data. Additional taxon sampling using GH gene sequences should resolve further relationships among various cypriniform lineages.

Howes (35, 36, 38–42, 44–46) provided the first attempts to subject the Cyprinidae to rigorous cladistic analysis. Howes questioned the monophyly of the Cyprinidae (47), in part due to never attempting an all-encompassing phylogenetic analysis of the family because of insufficient taxonomic materials. He did, however, propose a diphyletic arrangement into two lineages, cyprinins and leuciscins, based on presence or absence of the anterior maxillary barbel with an accompanying maxillary foramen. In addition, Howes (47) provided a synopsis of those subfamilies he considered valid with their diagnostic characters. Howes’ many contributions to cyprinid systematics will be invaluable in determining morphological transformation series for our phylogeny of Cypriniformes.

Cavender and Coburn (8) provided the first published phylogenetic analysis of the Cyprinidae, providing ten morphological characters to diagnose the family. They also recognized two major lineages within the family: Cyprininae, containing three groups (barbins, cyprinins, and labeonins) and Leuciscinae, containing eight groups (tincins, rasborins, gobionin, acheilognathins, xenocyprins, cultrins, leuciscins, and phoxinins). These groups are similar to those proposed by Chen et al. (10). Relationships within and among some groups had considerably more character support than was found for other groups; for example, the placement of Tinca as the basal-most lineage within the Leuciscinae was considered putative due to a limited amount of available material. Cavender and Coburn noted that several groups (e.g., barbins, labeonins) require further diagnoses. Similar to Howes, the contribution of Cavender and Coburn to cyprinid systematics is invaluable and will provide basis for developing and examining additional morphological transformation series used in our phylogenetic analyses.

studies contain a number of analytical caveats, however, including irreversibility of character evolution (Sawada) and limited number of transformation series examined (Siebert).
will resolve issues associated with terminology and homology assessment of key morphological characters. We will sequence complete mitochondrial genomes of the type species from all 350 genera described species (whenever possible, adding new species as they are described) including representatives from each genus (see letters from Miya, Saitoh, and Zardoya). We will sequence three mitochondrial genes and four nuclear genes for at least 1000 species, representing over 30% of the currently recognized diversity. These data will allow us to answer the following major research questions. 1) Are the major cypriniform clades monophyletic? There are few characters that support the major groups in the Cypriniformes. In particular, the large family Cyprinidae, containing most cypriniform diversity, has few characters supporting its monophyly. 2) How do characters of phylogenetic significance develop from protein-gene interactions to fully formed adult morphology. 3) Are there constraints imposed on morphology that generate homoplastic convergence which inaccurately reflect underlying relationships? 4) How does polyploidy impact morphological diversity? 5) What is the age of cypriniformes and major groups within this clade. 6) What is the relationship between the phylogenetic history of this ancient lineage and known Earth history?

The central objectives of the Cypriniformes Initiative are to: (1) develop an interactive Web Portal (Reviewers – please visit research website to demo our product at http://museum.tulane.edu/cypportal) for synergistic research and educational activities on Cypriniformes (online keys; identification methods for otoliths – bony inner ear elements often found in archeological sites), education on diversity and distributions, cultural history of Cypriniformes, impacts of invasive species, and links to other sites like FishNet, Tree of Life, FishBase, etc. (please visit our general website at http://bio.slu.edu/mayden/cypriniformes/home.html); (2) reconstruct relationships of ca. 350 genera and species using entire mitochondrial sequences, nDNA, and a suite of morphological characters; (3) reconstruct relationships of 1,000 species using three mitochondrial genes and four nuclear genes and a suite of morphological characters; (4) examine development of 30 species, representing major clades, providing an essential framework for evolutionary, development, and “systems” biology questions in ongoing research programs involving zebrafish, Danio rerio (http://www.zfin.org; see supporting letter #25 from Zfin organization); (5) conduct inventories and rapid bio-assessments of remote aquatic ecosystems containing cypriniform species (see participant letters #2-24 and Page letter #1 from All Catfish Species Inventory); (6) describe species and produce classifications; (7) examine historical biogeography of cypriniform fishes from global to regional scales using recent and fossil taxa and their hypothesized relationships; (8) test hypotheses of molecular/morphological evolution by incorporating fossils and tectonic history of Earth; (9) provide online database for distributional studies, in collaboration with FishNet (http://speciesanalyset.net/fishnet; again, see our research site at http://museum.tulane.edu/cypportal), containing vital information on the changing distributions and impacts of invasive cypriniform species; (10) provide public information on diversity of cypriniform fishes, in collaboration with FishBase (http://www.fishbase.org), and their cultural and economic importance (see Web Portal and letter #30 from Dr. Rainer Froese); (11) provide an online key to the major groups and some species, especially commercially important and invasive species; (12) provide fundamental educational site on cultural importance, impacts of exotics, natural diversity, etc. (http://bio.slu.edu/mayden/cypriniformes/home.html); and (13) maintain the Tree of Life web pages for the Cypriniformes (http://tolweb.org/tree?group=Ostariophysi&contgroup=Telestei).

IV. Research Foci and Data Collection

A. Taxon Sampling. – With over 3,285 described species and over 2,000 undescribed species it would not be possible to exhaustively sample the diversity present within Cypriniformes. Given this circumstance we have devised a taxonomic sampling strategy encompassing known diversity and that will provide a substantial framework for current and future studies. Our outgroup sampling will include taxa from Characiformes, Siluriformes, Gonorrhynchiformes, and Clupeiformes. See Web Portal for complete listing of targeted taxa/outgroups (fossil and Recent) http://bio.slu.edu/mayden/cypriniformes/home.html.

B. Working Groups. – Research foci have been focused around the following three working groups. PIs associated with these groups are given in the ‘Working Groups’ section of in the Management Plan.
1. Development Working Group. – This group will focus its attention on a set of 25 ingroup genera (Cyprinus, Barbodes, Barbus, Labeo, Puntius, Danio, Tinca, Leuciscus, Pimephales, Myxocyprinus, Carpiodes, Ictiobus, Catostomus, Moxostoma/Hypentelium/Minytrema, Cycleptus, Nemacheilus, Schistura, Crossostoma, Homaloptera, Hemimyzon, Gastromyzon, Botia, Leptobotia, Cobitis, and Misgurnus) and three outgroup genera, Chanos (Gonorhynchiformes), Ictalurus (Siluriformes), and Astyanax (Characiformes). We will rear most taxa for development at Saint Louis University; others will be spawned and gathered from commercial trade. Complete developmental series will be established for the full set of taxa in order that homologous elements can be assigned, ambiguities in terminology can be resolved, and a complete understanding of developmental patterns within Cypriniformes can be developed. In addition, with the intentional incorporation of zebrafish into the developmental framework we have a substantial opportunity to contribute significantly to the functionality of the zebrafish as a model organism for addressing evolutionary questions (see letter from zebrafish community #25).

2. Molecular Working Group. – This group will focus its efforts on a set of 1,000 ingroup species distributed over the known diversity within this family. The breakdown for developing molecular character sets for cypriniformes follows a tree-based hierarchy as below in Molecular Character Sets. The mitogenomics labs (letters #18 & 21) will provide structure to the basal nodes of the tree by focusing their efforts on obtaining complete mitochondrial genomes for the type species (where possible) from each of ca. 350 genera inclusive of outgroups. The nuclear and mitochondrial gene labs will develop complete DNA sequences from matched genomic samples derived from individual specimens for the set of genes proposed. While all monotypic genera will obviously be sequenced (whenever possible), the exact distribution of remaining taxa will follow an algorithm allowing a uniform distribution of samples proportional to the number of individuals contained in each genus. A complete list of the genera to be sampled for the gene-sequencing portion of the proposal can be viewed on the Web Portal.

3. Morphological Working Group. – This group will focus its efforts on developing a complete matrix for the same set of 1,000 ingroup species of extant cypriniform species being studied by the molecular group. In addition to extant species, all significant fossil specimens will also be examined in order that transformation series (TS) can be coded in a consistent manner across all known diversity (see Web Portal for fossil taxa to be examined). Within this group, many of the PIs and associate researchers have long standing research programs centered on understanding the morphological diversity present within various portions of Cypriniformes. The primary challenge facing these researchers is to reconcile issues of homology for character states within TS across the entire diversity of cypriniformes, eliminate inconsistencies in terminology across groups, and obtain data for understudied or incompletely studied species for all TS. It is anticipated that the resultant matrix will contain complete morphological data for all 1,000 specimens and significant fossil taxa. It is anticipated that approximately 750 morphological transformation series will be coded for Cypriniformes. In order to complete the morphological data matrix within the time frame of the study we intend to evaluate at least 150 TS per year in the first through fourth years of the study, leaving year five free to review the resulting matrix, fill in missing cells for difficult to obtain taxa and analysis. In order to assure homology among characters within TS various anatomical and ‘homologous’ regions have been divided across PIs with each region being the primary responsibility of one or more PIs (see Table 1). In order to ensure that characters are unambiguous and can be interpreted by future researchers wishing to make use of the data assembled, each member of the morphological working group will discuss characters and codings at the annual ASIH meetings; all data and images of characters will be permanently deposited on the Web Portal for public use following the completion of this study.

C. Morphological Character Sets. – As with most taxa, morphology (usually osteology) has served as the fundamental character set for taxonomic and systematic studies on cypriniformes. This same suite of character types will feature prominently in our large-scale revisionary studies as well. The PIs assembled for this portion of our investigation have internationally renowned reputations for their outstanding contributions not only with cypriniform systematics but also in phylogenetic coding methods (59, 60, 65).

Table 1 – Classification of skeletal regions for zebra danio, Danio rerio, a model organism of Cypriniformes.
osteology, and development of these fishes (see Management Plan). Classical systematic studies on cypriniform fishes range from those focusing on single species or small species groups (3, 14, 28, 35, 64) to large-scale revisionary investigations (19, 37, 46, 65, 79, 83, 86). These studies and many more have all surveyed the osteological features of the cypriniform skeleton for a variety of character information, including such things as: consistent variation in numbers and shapes of bones, interrelationships of bones, modes and patterns of bone formation (endochondral/dermal/membrane), tendon and ligament connections (and associated osteological formations), placement of foramina, unique bony elements, scale anatomy, otolith anatomy, lateral and cephalic sensory canal shape and pore development, modification of fin rays into spine-like elements, modifications of the Weberian apparatus, and scale morphology. When carefully examined for consistent morphological differences and similarities for these general qualities of bones and muscles, many character TS are possible. In our analysis, we fully anticipate deriving over 750 character TS from the skeleton. This is entirely reasonable given that PI Mayden (65) provided 362 TS (>400 taxa) and PI Coburn (8, 13) provided 163 largely different TS (>250 taxa). Morphological characters have demonstrated their utility in resolving phylogenetic relationships in these taxa and many others at various taxonomic levels.

Our analysis of the osteological features of Cypriniformes has been developed to maximize efficiency and consistency while at the same time minimize confusion as to the TS being used for phylogenetic analyses. Like the molecular portion of our study, we have divided the “character rich elements of the skeleton” among PIs for character elaboration. With the large number of taxa being examined (>1,000 species) it is important to stay focused on particular anatomical elements for consistency of hypotheses of homology and TS development. Table 1 identifies regions of the skeleton used by the zebrafish community for developmental studies and used in our investigation, and identifies the PI responsible for an inventory of characters within each region. All PIs will work with both cleared and stained and dry skeletal preparations (articulated and disarticulated) and will share specimens and/or skeletal elements.

### C. Molecular Character Sets.

This proposal includes three major molecular data sets, selected mitochondrial genes (cyt b, 12S, 16S), complete mitochondrial genomes, and selected nuclear loci (S7, Rag1, Growth Hormone (GH), Rhodopsin (Rh)).

#### 1. Genome Sampling.

We will use a root, stem, and leaf approach as our strategy for examining molecular evolution in cypriniform fishes. The root approach will use a smaller subset of cypriniform taxa to examine phylogenetic relationships within this large and diverse clade using complete mitochondrial genome sequences (ca. 16-17 kb). Several recent studies have demonstrated the phylogenetic usefulness of the mitochondrial genome for resolving deep phylogenetic divergences in fishes (50, 67, 69, 82). In addition to the large amounts of new sequence data, the order of individual genes provides another source of characters (62, 68). To determine basal-relationships among Cypriniformes, we will sequence ca. 350 specimens, representing the type species of all cypriniform genera, whenever possible, plus outgroup taxa. The mitogenomics laboratories (see Management Plan & letters # 18 &21) will be responsible for the sequencing of these specimens.

Mitochondrial genes have not always provided resolution at intermediate (i.e., stem) levels of phylogenetic divergence (30, 80, 89); nuclear genes have, however, demonstrated their utility in resolving
relationships across broad phylogenetic levels (i.e., root, stem, and leaf) (12, 73, 74, 102). In addition, several mitochondrial genes (e.g., cyt b and 12S & 16S rRNA) have been used extensively to resolve species-level relationships (e.g., 30, 80, 89). The stem and leaf approach will use a combination of nuclear (S7, RAG1, GH, Rh) and mitochondrial (cyt b and 12S & 16S rRNA) genes. Eight laboratories will be conducting the gene sequencing (see Management Plan).

2. Mitochondrial Genes. – Mitochondrial gene sequences have been widely used to recover relationships of cypriniform fishes. The vast majority of these focus on cyprinids (primarily European and North American taxa), very few include catostomids, balitorids, cobitids, or gyrinocheilids. Several mitochondrial genes have been used repeatedly to resolve phylogenies of cypriniform fishes.

Cytochrome b is the workhorse of molecular phylogenetics in vertebrates and as expected is the most widely used gene for cypriniform systematics. Lydeard & Roe (54) reviewed the utility of cyt b for assessing relationships in actinopterygian fishes and argued, with some caveats, that this gene had the potential to resolve relationships at a wide range of taxonomic levels. This gene has been used in interspecific studies among closely related taxa (7, 30, 84, 88), and also to recover phylogenetic relationships of most cyprinid diversity (15). It is widely used because it is a protein coding gene and thus knowledge of rates or probability of changes at each codon position can be accommodated in phylogenetic analyses. When used for higher-level phylogenetic problems a consistent pattern emerges. Typically trees based on cyt b are resolved at the base and at the tips but interior nodes often lack resolution.

The two ribosomal RNA genes, 12S and 16S, code for the small and large subunits of the mitochondrial ribosome. These have been used primarily in analyses of species groups and higher taxa (17, 29, 89). These tend to more conserved than protein coding genes and often exhibit little divergence between closely related species; however, Meyer et al. (66) used 16S rRNA to investigate relationships among species within the genus Danio and Okazaki et al. (72) used 12S rRNA for species within the genus Rhodeus. The use of rRNA genes is complicated because of the potential for indels, which makes alignment difficult.

3. Complete Mitochondrial Genomes. – Complete mitochondrial genomes have been used for reconstructing phylogenetic relationships of fishes (67). Saitoh et al. (82) published a phylogeny of Ostariophysi based on complete mtDNA sequences but only included seven cypriniformes from three families. Collaborators Miya, Saitoh, and Zardoya (see accompanying letters) are working on cypriniformes and are committed to the mitogenomics for this project, sharing all information, to aid in the international effort.

4. Nuclear Genes. – Single copy (presumably independent), protein-coding nuclear genes represent a good alternative to mitochondrial genes for estimating deep divergences and avoiding pitfalls associated with use of only mitochondrial genes. A suite of these genes has been selected for use in our molecular systematic studies in Order Cypriniformes. Candidate nuclear loci were selected based on the availability of nucleotide (nt) sequence data for cypriniform and related taxa. Most of the candidate loci code for proteins with a minimum of 200-300 amino acid (aa) residues. A recent analysis of GH gene sequences (12) by PI Bart, suggests that this length provides sufficient numbers of nt and/or aa characters to resolve phylogenetic relationships of major cypriniform taxa. The candidate nuclear genes show varying degrees of sequence divergence, increasing the likelihood that they will provide phylogenetically informative characters at several taxonomic levels within the Order Cypriniformes. In selecting candidate loci, consideration was also given to ease of isolation and amplification (i.e., need for obtaining messenger RNA and primary cDNA sequence data for designing amplification primers) for major cypriniform lineages (successfully done in PI Bart’s laboratory and being done in others).

The S7 ribosomal protein is a single-copy nuclear gene that is divided into seven exons separated by six introns (1, 2, 9). Fish primers for S7 were published by Chow et al. (11). The first intron is approximately 700 bp in length and has been used successfully in studies of fish phylogenetics. Lavoue et al. (52) used the first and second introns of S7 to resolve relationships of African electric fishes. These data provided a high degree of resolution and they observed no evidence of saturation. Wang et al. (102) use S7 to resolve relationships of 16 species of Asian cyprinid fishes with high bootstrap support across

Tree of Life: Systematics of Cypriniformes

Page 9

Project Description
most of the tree. They observed pairwise sequence distances of 0.0232 to 0.6659 across the included taxa. PIs Shunping and Mayden (32) are currently collaborating and have a manuscript on Cypriniformes and S7 demonstrating the excellent utility of this gene submitted for publication.

Growth hormone (GH) is a single chain, pituitary specific hormone that is essential for promotion and maintenance of somatic growth in vertebrates. In cypriniforms, GH comprises 5 exons, and 4 introns coding for a protein of approximately 220 aa residues. A number of GH sequences are available for cypriniforms and other fishes, and studies have shown that it is an excellent locus for phylogenetic inference (reviewed in 12). A recent phylogenetic analysis from PI Bart’s lab resolved basal relationships among cypriniform taxa with high bootstrap support (12). The success of GH in resolving basal cypriniform nodes also suggests that it is a useful benchmark for gauging the resolving power of other nuclear loci proposed for use in this study. GH nt sequence divergence ranges 0.03% to 12.4% (mean=7.2%) among 11 old world cyprinid fishes. This group is on average 15.1% divergent from the loaches (Cobitidae) of the genus Misgurnus, and 13.8% divergent from the catostomid, Ictiobus bubalus. Despite divergences ranging from 12-15% among cypriniform families, GH shows little evidence for saturation or loss of phylogenetic signal, even at the third codon position.

Another attractive locus is rhodopsin (Rh), expressed in the vertebrate eye and codes for light absorbing pigment. Rh codes for a 350 aa long protein, lacks introns (22), and has been sequenced for a variety of fishes. Among Cypriniformes, Rh data are available only for Cyprinus carpio, Carassius auratus, and Danio rerio (all Cyprinidae). Rh sequence comparisons among these taxa suggest that this locus is slightly more variable than GH, with divergences ranging 6-17% (mean=12.8%). By comparison, GH sequence divergences among these taxa average 9.7%. Mean Rh sequence divergence between these minnows and other otocephalans (Astyanax, Chanos, and Engraulis) is 20.4%, similar to the divergence seen in GH when minnows are compared to five siluriform taxa (21.0%). Among the three cypriniform taxa, 232 of 1068 total aligned positions (including gaps) are variable. Substitution patterns are as follows 65% 3rd pos, 23% 1st pos, and 12% 2nd pos. Average aa sequence divergence for the cypriniform taxa is 7.6%. Thus, Rh will be a useful marker for phylogenetics of Cypriniformes.

The recombination activating gene (RAG 1) is an endonuclease catalyzing the site specific V(D)J recombination process occurring during development of T- and B-lymphocytes (77). RAG genes are expressed mainly in teleost kidney (75). RAG 1 is a large locus of approximately 1300 aa in fishes and comprises three fairly long exons and two comparatively short introns. Sequence comparison of an 1100 nt fragment from the 3′ end of the gene among three cyprinid taxa (Danio, Carassius, Pimephales) suggests that RAG 1 is intermediate in sequence divergence to GH and Rh. Sequence divergence between Danio rerio and Carassius is 15.2%. Mean sequence divergence between minnows and other ostariophysans (Amieurus and Corydoras) is 23.0%, similar to GH and Rh divergences among ostariophysans. There are a total of 204 substitutions among cypriniform taxa with most substitutions at the third codon position (65%) followed by first (18%) and second (17%) position substitutions. Average aa divergence within these Cypriniformes is 10.1%. RAG 1 has been successfully used to resolve otocephalan and protoacanthopterygian relationships (106). The relatively long exons, availability of amplification primers, and sequence divergences suggest that RAG 1 should be used in this study. This gene is currently being examined in various PI labs (Bart, Harris, Mayden, Simons, Wood).

V. Data Analyses

A. Alpha-level taxonomy. – In addition to the 3,285 described species of Cypriniformes there exist over 2000 new species awaiting description. Field work done by the All Catfish Species Inventory (see Page letter, #1) and the present PIs or their collaborators in undercollected portions of the Cypriniformes distribution are likely to turn up additional diversity. Members of the Cypriniformes AToL subscribe to the Evolutionary Species Concept as our paradigm for diversity at the alpha level and utilize a variety of surrogate concepts to operationalize this for practical application. Among these are the phylogenetic species concept as defined by Wheeler and Platnick (76) wherein a species is described as "the smallest aggregation of (sexual) populations or (asexual) lineages diagnosable by a unique combination of character states." Descriptions of new species (phylogenetic placement based on other data) will be
grounded largely in standard methodologies of examining external and internal morphologies, truss measurements, fin-ray and scale counts, tooth characteristics, and coloration patterns. All descriptions will consist of standard diagnoses, type descriptions and determination, comparative analyses, suitable statistical analyses (PCA, sheared PCA, ANOVA, etc.), tabulation of information, and figures consisting of line-drawings, distribution maps, and black and white and color plates. Several of the PIs on this proposal and our foreign collaborators are heavily involved in alpha-level taxonomic studies of Cypriniformes and have described species; these PIs and their collaborators not on this proposal will describe species in during the course of this investigation (see our Web Portal for listing of the numerous ongoing studies by our group). PIs Kottelat, Mayden, Wood, and Harris all serve as editors of internationally recognized publications wherein species descriptions, revisionary studies, and systematic analyses are published, and are familiar with the necessary and sufficient elements of species descriptions. Professor Kottelat will serve as the editorial advisor for all species descriptions; he is managing editor of Ichthyological Explorations of Freshwaters and has published many taxonomic studies and inventories, and described many species.

Holotypes of new species collected by members of the Cypriniformes AToL Project will be deposited in an appropriate institution in the country of origin, if such an institution exists. If no such institution exists in that country, the holotype will go to an appropriate institution on the same continent or be kept at a PI institution where it would be transferred upon request of the appropriate official of the country of origin, provided an appropriate repository is created. When an existing museum specimen is designated the holotype of a new species, the responsible curator of that collection will determine that specimen's deposition. Paratypes and other specimens collected during the project will be broadly distributed as agreed to by the author(s) of the new species and the PIs. At least one paratype of each new species described with funding from NSF will go to a U.S. institution and one will go to an institution in the country (or continent) of origin.

B. Matrix Construction and Phylogenetic Analysis.

1. Molecular Data. – Sequence alignment is a highly controversial issue in contemporary systematics. Problems arise with complexities associated with choice of optimality criteria and sensitivity (76). Given that alignment is such dynamic and rapidly changing part of systematic biology, issues and controversies are likely change dramatically over the period of this grant. For some protein coding genes, such as cyt b, alignment issues are trivial. For others, such as rRNA genes, with large numbers of indels, alignments can be exceedingly difficult. We will use secondary structure models as an aid to alignment of rRNA genes (16, 101). Alignments will be run using Clustal X (51, 99). Large, complex datasets will be aligned on the University of Minnesota supercomputer system. Increased taxon sampling ameliorates many alignment issues (89). We also anticipate that regions that cannot be unequivocally aligned will be excluded from analyses.

Phylogenetic analyses will be run using the latest versions of PAUP* (98) and MrBayes (49). We will run parsimony, likelihood, and Bayesian analyses. Support in Parsimony analyses will be based on nonparametric bootstraps (21) and Bremer support (6). We will run separate and combined analyses of molecular data. Likelihood analyses will be used to identify appropriate evolutionary models (48, 78). Model based approaches are increasing in their complexity and it is now possible to implement multiple models in a single analysis using MrBayes (partitioned Bayesian analysis). We will explore the use of partitioned analyses with partitions both within and between genes. Bayesian analyses will be run for at least 2 million generations. The distribution of likelihood scores will be examined for stationarity and trees generated prior to achieving stationarity will be discarded. Remaining trees will be summarized in a majority role consensus tree; the frequency with which clades are recovered is interpreted as the posterior probability of that clade. The use and interpretation of these scores has been questioned (97); however, we point out that this is the case will virtually all other measure of support in phylogenetic analyses (34). Combination of datasets is a controversial topic and different genes may exhibit different histories as a result of paralogy, incomplete lineage sorting or lateral transfer (63). We will test for combinability of
genes using tests such as the incongruence-length-difference test (20). Unless there are strong reasons not to combine the data we will rely on combined analyses (103).

2. Morphological Data. – Concatenation of Recent and fossil osteologies of cypriniformes will be accomplished through the PIs sharing specimens and images both through the portal and at meetings. Currently, PIs Mayden and Arratia and foreign collaborators Chang and Huangzhang are examining both fossil and Recent taxa for character homologies, descriptions, and systematic studies. Data from Recent and developmental osteologies of Cypriniformes will similarly be accomplished. PIs Arratia, Coburn, and Mabee routinely employ developmental characters and homologies in anatomical and systematic studies of these fishes, and they will focus their efforts on skeletal characters of particular homology concern or of special phylogenetic interest. As a consequence, it is critical for them to maintain close contact with morphologists working on adult osteology as well as with each other. This will be accomplished by meeting with each other once a year in addition to ASIH and the large group meeting. Developmental data will be in a form that can be integrated into a larger morphological data matrix to either supplement previously identified character states or to provide new character states.

A large literature exists on analysis of morphological data (reviewed by 104). Many members of this collaboration are recognized for their expertise in the collection and analysis of morphological data and are aware of the complexities and possible problems. Morphological data will be analyzed with and without fossil data to examine the effects of incomplete taxa using parsimony, likelihood, and Bayesian methods as described above.

3. Combined dataset. – All data (molecular and morphological) will be combined in a large megamatrix for final analysis. The megamatrix will be analyzed using parsimony, likelihood, and Bayesian methods as described above using an evolutionary model suitable for morphological data (53). Methods for analysis of complex datasets are developing rapidly and we note that the field will have changed dramatically by the time the megamatrix is complete. In particular we anticipate that development of complex Bayesian models capable of incorporating multiple genes and morphology will continue (e.g., see 71).

The incorporation of fossil data, together with extant morphological and molecular data, raises the exiting possibility of developing a large phylogenetic hypothesis with multiple calibration points. Given the interesting distribution of Cypriniformes (Laurasian with presumed invasions of India and Africa) we will be able place approximate dates on origin of major clades, vicariance events, and dispersals.

C. Data Management. – The study of cypriniform systematics and biodiversity will result in a great diversity and number of data, datasets, and other information that will need to be archived and available for use by the collaborators, and ultimately to the public. The Web Portal discussed above will allow collaborators to contribute individual DNA sequences, NEXUS formatted files for molecular and morphological datasets, image files of specimens and anatomical characters, phylogenetic trees and taxonomic classifications. These data will result from ongoing research and will be derived from historical studies to compile a complete accounting of the systematic and taxonomic diversity of the group. Various customized online forms will be available to collaborators to upload information into the project database. Please visit our demo Web Portal for researchers at http://museum.tulane.edu/cypportal to view a demo of some of the functional elements of the site.

VI. Feasibility

As addressed below, final product of our research initiative will be major papers and web contributions to the phylogeny of these important fishes, inclusive of major and extensive morphological data sets for Recent and fossil taxa, the placement of important fossils (taxa listed on Web Portal). Together, the phylogenetic trees based on morphology and molecular analyses, combined with data from fossils will result in estimates of ancestral forms and times of divergences for major groups.

We have designed this study such that the data and analysis proposed is unquestionably attainable within the 5-year time frame. PIs have already begun (and have been doing for sometime) developing data management, image linkage, data input and “verification/curation of informatics” for placement on the Web Portal. Current taxon limits on tree searching algorithms can accommodate large data sets (up to
ca. 4,500 taxa). PAUP* and associated programs needed in analyses are available to PIs and adequate for our analyses. However, given the rate of software development other options will likely be available.

Our megamatrix will be established using a community-level basis. Multiple PIs and collaborators on this project are active researchers and publisher and have all agreed to our community-developed **Cypriniformes Team Charter** for sharing and collaborative research and publication (see supplementary documentation and attached letters of support). Because of the involvement of a broad array of collaborators outside of the core participants, completion of the matrix is quite feasible. The plan to hold a series of workshops to bring the collaborators together to assemble the matrix facilitates completion within the 5-year time-frame of this proposal.

Finally, we have a very close association with the **All Catfish Species Inventory** group wherein our members will receive materials of Cypriniformes and we will provide their group with Siluriformes. This will enhance the feasibility of our obtaining materials. Furthermore, through this initiative and our foreign collaborators and the international-level of work by the PIs we anticipate no problems with collecting permits. None of the taxa occur natively in Brazil and Australia, two countries most difficulty to export materials; India is also a difficult region but L. M. Page is negotiating with their government with the **All Catfish Species Inventory** and we have PIs that sample in India and students from India involved in our team.

**VII. Expected Results and Dissemination of Knowledge**

A. **Symposia.** – PI Mayden has arranged for an initial meeting and an approved symposium for participants (and other interested persons) at the 2004 European Congress of Ichthyology XI entitled “**Systematics and Biodiversity of Cypriniformes**” (see symposium announcement notice #34 in supplementary documents). PIs Arratia and Mayden are currently pursing (with proposal to Oxford Press and Academic Press) to publish these papers. PI Mayden has also been invited to discuss this international project to request international participation and funding of different aspects of the project in Beijing at the International Congress of Zoology (see notice #35).

B. **Publications.** – PIs Mayden and Arratia have already been approved for a contract to publish the entire compendium of this research effort as a book by the same name through Pfeil Publishing (see letter #33). Many other papers will result from this multi-institutional and PI effort. The members of our Cypriniformes Team are all active publishers and have unanimously agreed to our Cypriniformes Team Charter (charter document listed above and letters #24) regarding sharing information and publications. Our members are actively involved in many publications with one another and other collaborators. A list of the numerous current projects and publications relating to cypriniformes by these PIs is provided at http://bio.slu.edu/mayden/cypriniformes/home.html. This collaborative effort has already galvanized many of our participants into new projects and this is fully expected to continue, especially with funding, in the future on focused efforts to achieve our goals in both the educational/outreach and research areas.

The PIs have rapid publication outlets and some are editors of journals. Example publication outlets appropriate for various aspects of our study include *ZooTaxa, Ichthyological Explorations of Freshwater, Copeia, J. Fish Biology, Systematics and Biodiversity, Japanese J. Ichthyology, Molecular Phylogenetics and Evolution, Systematic Biology, Cladistics, Zoologica Scripta,* and *J. Biogeography.*

C. **Electronic Dissemination.**

1. **Web Portal.** – A Web Portal has been developed to facilitate collaboration among researchers studying the phylogeny of cypriniform fishes (two current sites active for viewing – our basic educational site at http://bio.slu.edu/mayden/cypriniformes/home.html and our primary research site at http://museum.tulane.edu/cypporal. After collaborators register to use the portal they are automatically placed on the project mailing list where many of the discussions of the various working groups will take place. All mailing list messages will be archived and accessible through the portal. The portal will also serve as a document repository for all project-related files and publications. The Web Portal will support user authentication, allowing collaborators full access while only providing access to published data for non-collaborators. Our webpage for the general Cypriniformes portal can be viewed at the web address above; a demo package of how the database will work for users is at the latter site listed above. For
example, users will deposit images of anatomy, characters, and specimens, sequences, NEXUS files, primers, reaction profiles, aligned sequences, and phylogenetic analyses, all for sharing across the working group and ultimately the public; one can upload and download information.

One of the primary functions of the Web Portal is to allow collaborators to query and display information contained within the database as well as data from online natural history databases containing cypriniform specimen records. The Portal allows users to search any of the fields contained within the project database and return all related records, images and files. The Portal will also support queries to FishNet, an online distributed network of natural history fish collection databases, to gather related information on available specimen records. By facilitating access to specimen museum records, users will be able to examine historical identification and distribution records of many cypriniform species. The Portal will include GIS capabilities enhancing geographical queries and visualization of species data from both local and distributed sources. The open source Mapserver developed by The University of Minnesota will be the basis of the portal web-mapping engine. By providing these tools the Web Portal will provide a single point of access for users to access and utilize worldwide cypriniform information.

2. Tree of Life (Maddison’s ToL Web project and initiative). – Our group will lead the effort on updating the tolweb site for Cypriniformes. PI Mayden has been asked by the Maddisons and Lundberg to develop this section of the tree. Currently there are no phylogenies for these fishes on ToL Web.

3. Dynamic Taxonomic Tree. – Collaborators will also use the portal to view and suggest changes to our current understanding of Cypriniform relationships. A webpage will be available within the research Portal allowing users to view a list of current names, their classification and references. Users wishing to make a change to the classification or nomenclature of an item in the list will have the option to click an item in the list and suggest a change. Each suggested change to an item will be recorded in the database and displayed along with that item. When a change is suggested an email will automatically be sent out to the collaborators mailing list notifying them of a suggested change in the taxonomic tree, sparking conversation regarding the change. The lead PI will receive an additional email with a link to make decisions on proposed changes. Acceptance of the change causes the item to be reclassified or renamed within the list. Previous classifications/nomenclature are archived to serve as information history.

4. Internet Sites. – We have developed two working sites for communication, our basic information location at http://bio.slu.edu/mayden/cypriniformes/home.html and our research Web Portal (ultimately accessible through the basic location) at http://museum.tulane.edu/cypportal (please visit for a tour).

5. Online Keys and Identifications. – We have developed online and illustrated keys to the major groups within the Cypriniformes and this process will continue during our study. We will have keys to all genera, illustrated by participant Dr. David A. Neely (working in PI Mayden’s laboratory; see his images on our interactive keys), and will also have keys developed for many of the species (see demos at Key location http://bio.slu.edu/mayden/cypriniformes/home.html) (genera Moxostoma (Catostomidae) and Cyprinella, Notropis (Cyprinidae)). Our identification section will also include online methods for identification otoliths, inner ear elements of these fishes that are difficult to ID and are often found in archeological sites and rat middens.

C. Public Workshops and Meetings. – Our working groups will communicate regularly via email and the discussions on the Web Portal (and possibly via video conferencing on Internet 2 when PIs get hooked up with hardware; I2 currently available at all institutions). We will have annual workshops on the project at annual ASIH meetings to discuss various aspects of the project, including taxon sampling, diversity discovered, manuscripts in progress, data management, character homology, analyses, matrix concatenation, and emerging ideas. This project had its origins basically out of the Deep Fin Initiative (www.deepfin.org) actively being promoted by Guillermo Orti and PI Mayden. PI Mayden will present at the 2004 ASIH meetings on Deep Fin and has arranged for a workshop regarding the importance of large-scale collaborative efforts in systematic ichthyology at these meetings. PI Mayden has arranged for an initial meeting and an approved symposium for participants (and other interested persons) at the 2004 European Congress of Ichthyology XI entitled “Systematics and Biodiversity of Cypriniformes” (see symposium announcement notice #34 in supplementary documents). PI Mayden will also discuss this
international project to invite international participation and funding of different aspects of the project in Beijing at the International Congress of Zoology (see notice #35). All of the PIs have active research programs and customarily travel within and outside of the US, permitting up-to-date communication that will also facilitated by our Web Portal.

D. Education

This research will result in training of numerous graduate and undergraduate students in systematic biology. Previous funded research by the PIs has resulted in significant training of underrepresented persons and we fully expect that this will continue in this collaboration. Saint Louis University is committed to providing three fully supported (salary and tuition; >$500,000 for this project; one RA will be an international students working on Cypriniformes from either Russia, China, or Japan.

Our project also includes the development of several major educational initiatives if this proposal is funded. The Aquarium of the Americas, New Orleans, in collaboration with PI Bart is committed to producing live and stationary exhibits on Cypriniformes and this project for public consumption. We have arranged for Dr. James F. Scott of Saint Louis University (see http://www.slu.edu/colleges/AS/ENG/faculty/scott.html for examples of his work), in collaboration with other SLU PIs, to develop and produce at documentary on this international AToL initiative, focusing on cultural issues with these fishes, their diversity, the international collaborative nature of modern-day research from the field work to the molecular laboratories, and how we go about reconstructing the Tree of Life. The anticipated title of this movie is “The Making of the Tree of Life for Our Earth’s Most Diverse Freshwater Fishes.” Dr. Scott has an experienced writer and producer of documentaries for PBS and NOVA (see letter #32). Our Web Portal will distribute massive amounts of information on Cypriniformes on issues relevant to diversity, taxonomy, distributions, cultural issues, online keys, identification sections for otoliths, various “media products” (movies, color images, sounds produced by these fishes, spawning), the impacts of exotic species on natives and the AToL project (see supporting letters #26-32 on these planned projects and the importance of our educational initiative on these fishes).

Finally, PI Simons will collaborate with the education and exhibits staff at the Bell Museum of Natural History, University of Minnesota to develop educational programming and exhibits reflecting this research. Minnesota issues more fishing licenses per capita than any other state in the nation and its citizens have a strong appreciation for the natural world. We will build on this interest, creating educational products that use local issues such as invasive Asian carp and local cypriniform diversity to introduce the public to global issues, the importance and relevance of research in systematics, and documentation of biodiversity. This will result in two major products: 1) an outreach program aimed at 3rd-6th grade students, and 2) a museum exhibit suitable for a national audience (letter #26).

Literature Cited


