

## Building a Centralized Database for Kentucky Fishes: Progress and Future Applications

Robert L. Hopkins II,<sup>1</sup> Michael D. Burns, Brooks M. Burr, and Lisa J. Hopman

Department of Zoology, Southern Illinois University, Carbondale, Illinois 62901

### ABSTRACT

Kentucky has a long history of ichthyological collection and study. Voucher specimens are available for collections dating back to 1870 and are housed in various museums and research collections in the United States. In this paper we present a descriptive overview and the current status of a project aimed at building a centralized database for vouchered records of Kentucky fishes. To date, we have entered over 51,000 records into the database. Each of these entries contains at least three vital pieces of information: species identification, georeferenced locality, and time of collection. With increasing recognition of the potential value of museum-based data in biological, ecological and conservation studies such a centralized database will serve as an important scientific resource for the study of Kentucky fishes.

KEY WORDS: fishes, database, Kentucky, biodiversity, historical, museum

### INTRODUCTION

Ichthyological investigations in Kentucky date back to the earliest scientific documentation of the state's fish fauna in the "*Ichthyologia Ohiensis*" (Rafinesque 1820). In this work, Rafinesque provided natural history and descriptive information for fishes mostly from the Ohio River near the falls at Louisville. Only a small number of collections are known prior to 1820 and include anecdotal accounts of larger species, e.g., from the Ohio River (Pearson and Krumholz 1984). Woolman (1892) conducted a more comprehensive survey of Kentucky fishes and documented statewide historical distributional information prior to widespread anthropogenic influence. More recent literature on Kentucky fishes includes a distributional catalogue (Evermann 1918), state fish book (Clay 1975), a species checklist (Burr 1980), an Ohio River status and distribution update (Pearson and Krumholz 1984) and a statewide distributional atlas (Burr and Warren 1986). The Kentucky State Nature Preserves Commission, the Kentucky Division of Water, and the Kentucky Department of Fish and Wildlife Resources regularly conduct surveys and produce reports which contribute to the understanding of fish distributions and the status of streams and rivers in the state. Other notable contributions were made by Minor E. Clark, William R. Turner, and those associated with the Ohio

River Valley Water Sanitation Committee who provided detailed descriptions of watersheds in the state, pre-impoundment studies, and detailed documentation of the Ohio River ichthyofauna, respectively.

The most comprehensive assessment of Kentucky fish distributions was provided by Burr and Warren (1986). They presented a detailed discussion of physiographic and hydrographic features in Kentucky, species accounts that included point locality range maps, narrative descriptions of distribution, systematics, habitat, and a recommended conservation status. The distributional data in the atlas span from 1818 to 1985 and include records from all major museum holdings in the United States as well as records from personal collecting efforts by the authors. The range maps were fashioned by hand and no computerized records were created during the project. Burr and Warren (1986) is still the principle reference for Kentucky fishes. However, taxonomic revisions, descriptions of new species, and numerous new collection records since 1986 warrant updating of the distributional data.

In recent years there has been increasing interest in museum-based informatics (Graham et al. 2004). Natural history collections offer a wealth of temporal and geographic data tied to animal and plant collections. Developments in computing and analytical tools, especially geospatial technologies, are making these data more relevant and useful. As a result, many of the records used in producing

<sup>1</sup> Corresponding author e-mail: rhopkins@siu.edu



the distributional atlas (Burr and Warren 1986) now have been electronically cataloged. These data are available by written request to collection managers or, in some cases, via searchable internet databases. Efforts are underway by various museums and institutions to digitize and link databases online – allowing users to perform a single query across multiple databases. One such project for fishes is Fishnet 2 ([www.fishnet2.net](http://www.fishnet2.net)). This website allows a user to search fish records across 29 institutional databases with a single query, and results can be exported in a variety of formats for subsequent processing. However, this system has limitations: 1) only a relatively small percentage of the data are georeferenced, 2) there is no consistent schemata to the exported data, and 3) taxonomy rarely is updated for historical records.

In light of the current usage and potential utility of museum-based information for the study, conservation, and management of Kentucky fishes, we decided to build a centralized database for the state. Because many conservation plans and research projects are implemented within state boundaries, such a database would provide historical and contemporary data that would be readily available for analyses. The database also would include information linking records with museums containing voucher specimen(s). The intention of this project is to enhance research and conservation efforts focused on Kentucky's ichthyofauna.

The objectives of this multi-year, ongoing project are to 1) obtain records of Kentucky fish collections from all major museum and institutional holdings, 2) computerize and georeference those records and merge the data into a single dataset with a standardized information schemata, 3) complete error checking and data cleaning, and 4) update taxonomic and nomenclatural information. We provide an overview of our progress and the current status of the database.

## MATERIALS AND METHODS

In 2005, Southern Illinois University at Carbondale (SIUC) acquired the ichthyology collection from the University of Louisville (UL). A major task was undertaken to curate, georeference, and computer catalogue the

estimated 13,000 lots from UL. In addition to the UL project, we also began curating, georeferencing, and computerizing backlogged records and collections from SIUC, Kentucky State Nature Preserves Commission, Kentucky Department of Fish and Wildlife Resources, and the Kentucky Division of Water. This initiative provided the impetus to expand our efforts to create a comprehensive and centralized database of Kentucky fish records. This database is hereafter referred to as the Kentucky Fishes Database (KFD).

We estimate that greater than 90% of the vouchered collection records for Kentucky fishes are located in the following museums: California Academy of Sciences (CAS), Cornell University (CU), Illinois Natural History Survey (INHS), University of Kansas (KU), National Museum of Natural History (USNM), Ohio State Museum of Biological Diversity (OSM), Southern Illinois University Carbondale (SIUC), University of Michigan Museum of Zoology (UMMZ), and Tulane University (TU). We requested or downloaded all Kentucky fish records from these museums. Many other smaller collections also contain valuable data including Morehead State University (MoSU) and Eastern Kentucky University (EKU). We acquired paper copies of select records from MoSU and incorporated them into the database. The research collection at Murray State University had earlier been transferred and catalogued into the SIUC system.

Records received from museums were in various processed states. For example, some records were fully georeferenced and needed only to be converted to decimal degree format and fitted to the KFD schemata – whereas others were not georeferenced and contained numerous taxonomic and geographical errors. Most of the datasets, however, contained good descriptive information on localities but were not georeferenced and had out-dated taxonomy.

For each record processed, we georeferenced the site in decimal degree format using the North American Datum of 1983, updated taxonomic and nomenclatural information, and standardized the data schemata. Formats for data and headers were specifically chosen for quick incorporation into a geographical



Table 1. Contributing museums and respective number of records provided. Records from SIUC include the ~13,000 lots incorporated from the University of Louisville.

Museum	Contributed Vouchered Records
California Academy of Sciences	246
Cornell University	1051
Illinois Natural History Survey	4100
University of Kansas	598
Morehead State University	1261
National Museum of Natural History	908
Ohio Museum of Biological Diversity	2108
Southern Illinois University Carbondale	36,724
Tulane University	875
University of Michigan Museum of Zoology	3342
TOTAL	51,213

information system (GIS), which played a key role in data editing and enrichment. Ultimately, each record contains three vital dimensions: specimen identification, georeferenced collection locality, and time of collection. All records in the KFD are vouchered records. The incorporation of only vouchered records was intentional and assures that each entry is backed by a physical specimen; this allows the verification of species identification, use in research projects, and increases data integrity.

## RESULTS AND DISCUSSION

In 2005, the SIUC collection contained about 10,000 computerized records for Kentucky. As of early 2008, after the incorporation of the UL collection and processing of backlogged collections from state agencies, that number exceeded 36,000. At the end of 2008, SIUC will have more than 40,000 records of Kentucky fishes dating from the 1920s. We have voucher records for every extant species in Kentucky. SIUC's records – including those incorporated from UL – account for about 70% of the records in the KFD.

As of early 2008, the KFD contained over 51,000 records (Table 1). Excluding SIUC, the largest contributors of Kentucky records were INHS and UMMZ with 4100 and 3100 contributed records, respectively. The CAS and the USNM contained fewer records, but housed the majority of the A.J. Woolman and P.H. Kirsch collections, which are of historical value. Ultimately, we expect the KFD to

exceed 60,000 records. We are currently processing additional records from the University of Florida, University of Tennessee, Academy of Natural Sciences of Philadelphia, University of Alabama, Field Museum of Natural History, and the Harvard Museum of Comparative Zoology. In addition, we continue to process records from state agencies in Kentucky and from our own collection efforts.

The KFD includes records dating from the 1870s. Woolman's collections represent the earliest comprehensive surveys included in the database. Fewer than 20 pre-Woolman records, from unnamed collectors, exist for the entire dataset. The spatiotemporal coverage of the KFD (Figure 1) is statewide for selected time periods but most comprehensive for the 1970–1990 decades. The majority of vouchered specimens were added after 1950 with most specimens collected since 1980. Despite over 100 years of collection efforts, there are still some regions of the state lacking adequate coverage. Historically, the Nolin River drainage has lacked basic survey work. Likewise, additional collection efforts in the lower Green River, Tug Fork, Mississippi River, and direct Ohio River tributaries would permit a more complete documentation of Kentucky's ichthyofauna.

A consistent data schemata is being applied to all KFD records (Table 2). We encourage additional data enriching by incorporation of comments from collector's field notes, published papers, and reports when available. Such ancillary information (e.g., general comments on habitat, unusual observations, collecting methods/techniques, information on species not vouchered, designation as type specimens, etc.) can greatly enrich the point data. To date, we have georeferenced all records acquired (Table 1), but have not completed the totality of editing and enriching the data. Updating taxonomy from historical collections has proven to be time consuming, and we have made only moderate progress. For newer collections and especially those housed at SIUC we have completed most of the taxonomic updating.

Nomenclatural and taxonomic issues are being addressed using several methods. For species with contrasting biogeographies we are able to clean the data within a spatial



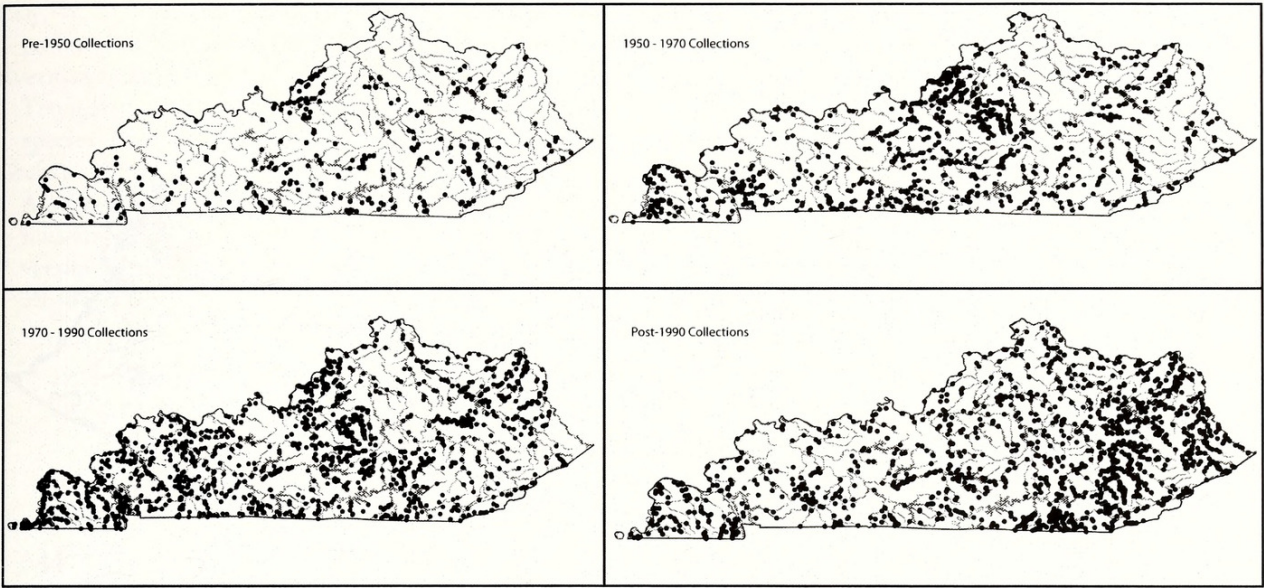


Figure 1. Overview by time periods of collection records currently contained in the Kentucky Fish Database.

format. For example, the *Etheostoma spectabile* (Agassiz), *E. simoterum* (Cope), *E. virgatum* (Jordan), and *Notropis rubellus* (Agassiz) complexes, each containing species with allopatric ranges are edited by plotting the target data within a GIS and assigning correct names to points using spatial-selection methods. In contrast, much editing of older names is being done in a standard database format. Despite extensive cleaning, we continue to find errors within the database. Based on past experiences, we estimate that perhaps 5–7% of the entries contain errors of some nature (e.g., out-dated taxonomy, misspellings, etc.) with 1–2% containing errors that seriously compromise the integrity of the data (e.g.,

incorrect locality information or misidentification). Focus on the three vital dimensions (identification, georeferenced localities, and time of collection) for each record is required to ensure high quality datasets.

The potential applications of the KFD data are multifold. Museum-based data can be integrated with other data types to address a series of questions ranging from conservation research to the study of ecological and evolutionary processes. Other applications include spatial analysis of biodiversity, the study and modeling of species distributions, analysis of range shifts, planning and developing future field projects, identification of potential areas for reintroductions, etc. To illustrate the utility of the KFD, we have completed a cursory spatiotemporal analysis of the distribution of the Trout-Perch (*Peropsis omiscomaycus* Walbaum) in Kentucky (Figure 2). Some localities for this species included in Burr and Warren (1986) are not shown on the map (Figure 2) because the KFD is restricted to vouchered records. In this case, the voucher for a collection of Trout-Perch Welter (1938) in North Fork Triplett Creek (Licking River drainage) reportedly deposited at MoSU could not be located.

The overall pattern that emerges (Figure 2) is the absence of Trout-Perch in recent collections from the western portions of its range. In contrast, eastern populations in Little Sandy River, Levisa Fork, Tygarts

Table 2. Data schemata for database with examples of included information.

Field Name	Information
Museum	SIUC
CatNum	64,942
Species	<i>Etheostoma derivativum</i>
Latitude	36.8822
Longitude	–87.1111
Waterbody	Whippoorwill Creek
Drainage	Red-Cumberland River
County	Todd
State	Kentucky
Day	19
Month	April
Year	2007
Collectors	RL Hopkins, LJ Fisk
Locality	at SR 106 bridge; 0.5 mi S of Claymour, KY
Notes	rifle area now inundated due to fallen tree



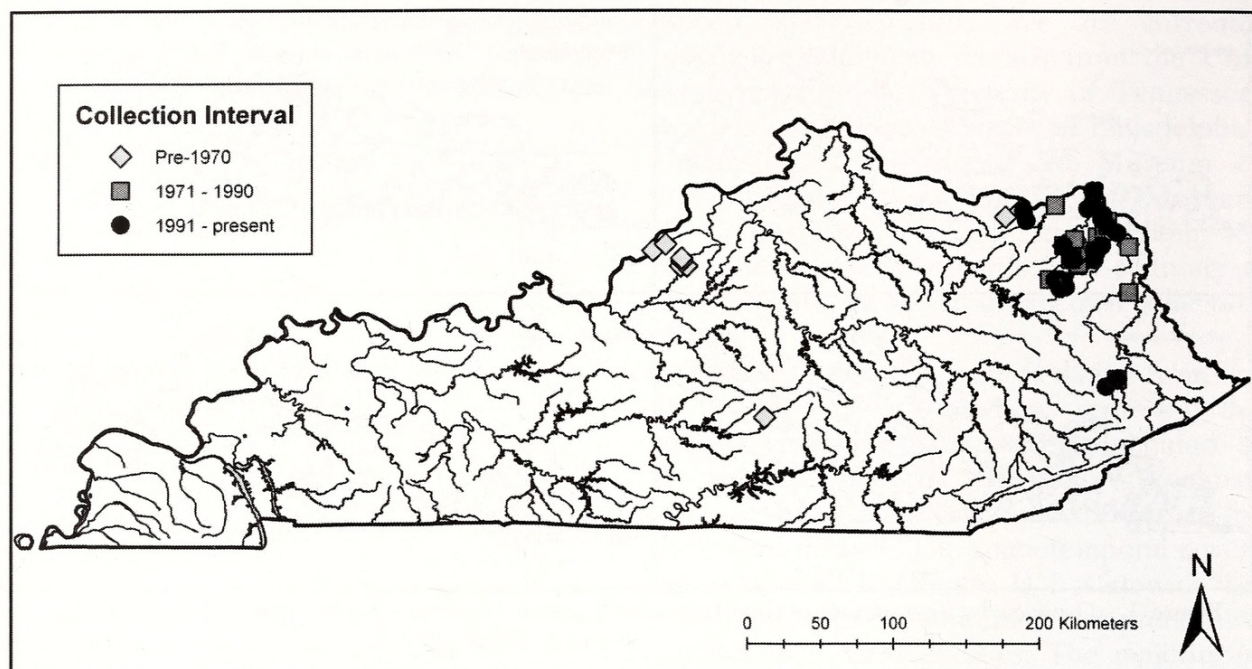


Figure 2. Spatiotemporal map of Trout-Perch (*Percopsis omiscomaycus*) records in Kentucky.

Creek, and Kinniconick Creek seem stable, although, no recent specimens have been captured from the East Fork of the Little Sandy River or Little Blaine Creek (Big Sandy River drainage). The apparent disappearance of this species in western areas warrants further research investigating the causes of its range contraction in Kentucky.

### CONCLUSIONS

Since building the database we have been able to respond to questions about the Burr and Warren maps much more effectively. We are not only able to provide locality information, but also the time of collection, holding museum, a list of other species captured at the site, etc. Upon request, we can produce maps and spreadsheets containing the aforementioned information. It is evident that such a database is a valuable tool for researchers and will increase the use of museum and research collections in ichthyological, ecological, and conservation studies.

Furthermore, as we begin to mine the data we are finding many noteworthy records for Kentucky fishes. We have evidence of range expansions and contractions and of new drainage records for several species. The collection from UL has provided a wealth of additional historical information for Kentucky fishes. In addition, there have been multiple

recent collections of rare fishes – an encouraging sign. We are preparing a manuscript providing an overview of some of the more noteworthy records stemming from the KFD project. We encourage museums of all sizes to georeference and computerize their fish collection records, so that they be more accessible for use in conservation planning and the study of evolutionary and environmental processes.

### ACKNOWLEDGEMENTS

We are indebted to the following museums and their collection managers for making these data available to us for processing: CAS, CU, INHS, KU, MoSU, USNM, OSM, TU, UL, and UMMZ. The following people devoted much time and energy into curation, georeferencing, and computerization of the records: K. Butera, A. L. DeFore, D. J. Eisenhour, S. L. Flood, R. R. Hopkins, and A. Richter. Partial support for this project was provided by the Kentucky Department of Fish and Wildlife Resources.

### LITERATURE CITED

- Burr, B. M. 1980. A distributional checklist of the fishes of Kentucky. *Brimleyana* 3:53–84.  
 Burr, B. M., and M. L. Warren, Jr. 1986. A distributional atlas of Kentucky fishes. Kentucky Nature Preserves Commission Scientific and Technical Series 4:1–398.



- Clay, W. M. 1975. The fishes of Kentucky. Kentucky Department of Fish and Wildlife Resources, Frankfort, KY.
- Evermann, B. W. 1918. The fishes of Kentucky and Tennessee: a distributional catalogue of the known species. *Bulletin of the Bureau of Fisheries* 35:295–368.
- Graham, C. H., S. Ferrier, F. Huettman, C. Moritz, and A. Townsend-Peterson. 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution* 19:497–503.
- Pearson, W. D., and L. A. Krumholz. 1984. Distribution and status of Ohio River fishes. ORNL/sub/79-7831/1, Oak Ridge National Laboratory, Oak Ridge, TN.
- Rafinesque, C. S. 1820. *Ichthyologia ohiensis*. W.G. Hunt, Lexington, KY.
- Welter, W. A. 1938. A list of the fishes of the Licking River drainage in eastern Kentucky. *Copeia* 1938:64–68.
- Woolman, A. J. 1892. Report of an examination of the rivers of Kentucky, with lists of the fishes obtained. *Bulletin U.S. Fish Commission* 10:249–288.



Hopkins, Robert L. et al. 2008. "Building a Centralized Database for Kentucky Fishes: Progress and Future Applications." *Journal of the Kentucky Academy of Science* 69(2), 164–169. <https://doi.org/10.3101/1098-7096-69.2.164>.

**View This Item Online:** <https://www.biodiversitylibrary.org/item/175664>

**DOI:** <https://doi.org/10.3101/1098-7096-69.2.164>

**Permalink:** <https://www.biodiversitylibrary.org/partpdf/335757>

**Holding Institution**

Smithsonian Libraries and Archives

**Sponsored by**

Biodiversity Heritage Library

**Copyright & Reuse**

Copyright Status: In Copyright. Digitized with the permission of the rights holder

Rights Holder: Kentucky Academy of Science

License: <http://creativecommons.org/licenses/by-nc-sa/3.0/>

Rights: <https://www.biodiversitylibrary.org/permissions/>

This document was created from content at the **Biodiversity Heritage Library**, the world's largest open access digital library for biodiversity literature and archives. Visit BHL at <https://www.biodiversitylibrary.org>.