Developing a Multi-species Anatomy Ontology for Teleost Fishes

Wasila Dahdul, Paula Mabee, *University of South Dakota;* **John Lundberg,** *Academy of Natural Sciences;* **Peter Midford,** *University of Kansas;*
Todd Vision, Hilmar Lapp, Jim Balhoff, Cartik Kothari, *US National Evolutionary Synthesis Center;*
Melissa Haendel, Monte Westerfield, *Zebrafish Information Network, University of Oregon*

Abstract

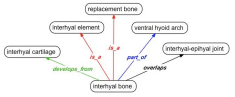
Anatomy ontologies are a practical mechanism to capture anatomical data in a systematic and computable manner. Multi-species anatomy ontologies in particular are an efficient way to represent the diversity of morphological structures in a clade of organisms, but they present challenges in their development relative to single species anatomy ontologies. The Teleost Anatomy Ontology (TAO) is a multi-species anatomy ontology for teleost fishes that began development in September 2007 by the Phenoscape Project (http://phenoscape.org) as a clone of the Zebrafish Anatomical Ontology (ZFA). It has since been extended by the ichthyological community, whose feedback is engaged by linking requests for new TAO terms, or changes of definitions or relationships, to an open mailing list that facilitates discussion of proposed terms. Since cloning from ZFA, over 400 new terms, primarily skeletal, have been added to the TAO. Phenoscape is using the TAO for annotation of evolutionary phenotypes from the comparative literature of the Ostariophysi, a large clade of approximately 9,000 teleosts that includes the zebrafish, minnows, catfishes, knifefishes, tetras, and the milkfish. Growth of TAO is focused on the skeletal system, because it is used most extensively as the basis of comparative morphology-based studies in ichthyology and paleontology. To facilitate interoperability with other anatomy ontologies, the upper level nodes of TAO are referenced to the Common Anatomy Reference Ontology (CARO) and terms in the TAO and ZFA are kept in synchrony; zebrafish terms are specified as subtypes (using the *is_a* relation) of teleost terms. We present several general solutions for the expansion of a single-species ontology core to accommodate multi-species diversity. These include refinement of definitions to represent the diversity of structures in many species and representation of taxon-specific relationships between terms.

Overview and Features of the Teleost Anatomy Ontology (TAO)

Structure and organization

The Teleost Anatomy Ontology (TAO) was developed by the Phenoscape project (www.phenoscape.org) in September 2007. Terms in the TAO are primarily structurally defined and upper-level terms are cross-referenced to the Common Anatomy Reference Ontology (CARO). We strive to maintain a single inheritance heirarchy in TAO, but some terms contain multiple *is_a* parents (below).

Relationships between terms



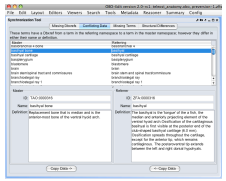
The primary relations used in the TAO are *is_a*, *part_of*, *develops_from*, and *overlaps*. *Overlaps* is used to represent the relationship between two entities that share a part; rather than creating a term for the shared part, the two entities are said to overlap each other. For example, interhyal bone *overlaps* interhyal-epihyal joint (above).

Cloning from zebrafish ontology

The TAO was cloned from the single-species Zebrafish Anatomy Ontology (ZFA), and zebrafish terms are subtypes of teleost terms (right). Many terms have subsequently been added to TAO which apply to teleosts but not to zebrafish (highlighted in green).

Synchronization between TAO and ZFA

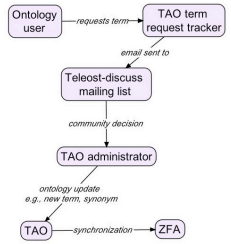
TAO and ZFA terms are cross-referenced and synchronized regularly. To automate this process, we have developed the **Synchronization Tool** (screenshot at right), a publicly-available plug-in for OBO-Edit2. The tool checks for missing cross-references between identically named terms, conflicting data between cross-referenced terms, missing terms, and structural differences such as differences in parent terms of cross-referenced terms.



Above: Using the Synchronization Tool in OBO-Edit2 to check for conflicting data between TAO and ZFA.

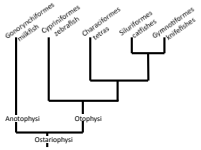
Community involvement in ontology development

The process of updating the TAO to include new terms, refined definitions, synonym additions, and structural changes begins with term requests from ontology users (e.g. people annotating images, phenotypes) through the Term Tracker request tracker. Requests trigger an automated email to the ichthyological community through the teleost-discuss mailing list. Discussion ensues, a decision is reached, and the request is summarized by the TAO administrator on the tracker page. The request is then closed, and the TAO is updated to include the requested change. Updates that also apply to the Zebrafish Anatomy Ontology (ZFA) are made during synchronization.



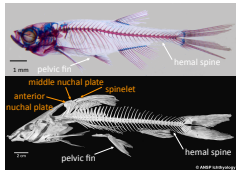
Anatomy for many fish species

The Teleost Anatomy Ontology (TAO) is a multi-species anatomy ontology for teleost fishes developed with the goal of representing the structural diversity present in this clade of fishes. The current focus of development is on the Ostariophysi, a diverse clade (>9,000 species) of freshwater fishes including the zebrafish, carps, tetras, knifefishes, and catfishes.



Above: Phylogeny of ostariophysan fishes modified from Fink and Fink (1981).

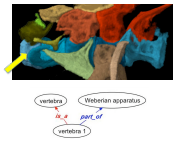
Development of the TAO is currently centered on the skeletal system, since it is often the focus of evolutionary studies in ichthyology and genetic studies in zebrafish.



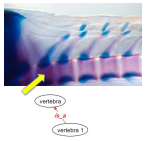
Above: Zebrafish and catfish skeletons. Many new terms (orange, above) are required in the TAO to represent the diversity of skeletal structures present in all teleost species.

Challenge: representation of taxonomically-variable relationships

Depending on the taxonomic context, a term may have different relationships with other terms. An example of this comes from the *Weberian apparatus* (below, left image), a complex of bones that are derived from the anterior vertebrae and associated structures that function in sound transmission from the swim bladder to the inner ear. The Weberian apparatus is a unique feature of otophysan fishes.



Above: *Vertebra 1* (yellow arrow) is *part_of* the *Weberian appartus* in zebrafish and otophysan fishes.



Above: Other teleost fishes lack the Weberian apparatus, hence *vertebra 1* is not *part_of Weberian apparatus* in these taxa.

We are exploring how to represent relationships with a taxonomic context by using entity post-compositions to assert taxon-specific anatomical relationships.

Challenge: creating definitions that apply to all teleosts

Term definitions inherited from the zebrafish ontology required generalization so that they apply to all teleost fishes. For example, the Weberian apparatus was defined as pertaining to the first four vertebrae in zebrafish, but it may include additional vertebrae in some teleosts. Thus the definition was changed to refer to "anteriormost vertebrae" to make it applicable to any teleost.

ZFA definition: The Weberian apparatus consists of the modified anteriormost four vertebrae and associated structures of otophysans which transduce sound or pressure waves from the tunica externa of the swimbladder to the inner ear (Alexander, 1962).

TAO definition: Anatomical cluster that consists of the anteriormost vertebrae and associated structures that connect the swim bladder to the inner ear.

Challenge: homology designation

TAO terms are structurally defined and homology is not implied between terms, nor is it implied if the same term is used in annotations for different species. Homology is a hypothesis that requires evidence and attribution for its designation.

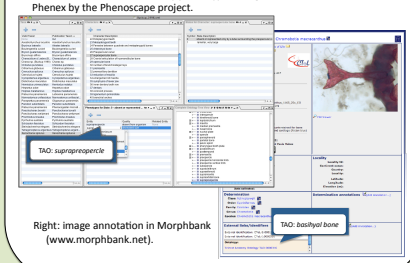
The Phenoscape project is using Phenote to record statements of homology which will be used in queries within the Phenoscape Knowlegebase:

Phenote homology table

Publication	Entity 1	Taxon 1	Entity 2	Taxon 2	Evidence
Britz and Hoffman 2006	supradorsal	Teleostei	claustrum bone	Otophysi	inferred from developmental similarity
Britz and Hoffman 2008	supradorsal	Teleostei	claustrum bone	Otophysi	inferred from positional similarity

Applications of the TAO

Below: annotation of evolutionary phenotypes using Phenex by the Phenoscape project.



Right: image annotation in Morphbank (www.morphbank.net).

Acknowledgements

We thank NSF DBI 0641025; NIH HG002659 and the National Evolutionary Synthesis Center (NESCent) NSF #EF-0423641 for funding. Thanks to K. Luckenbill forthe zebrafish MicroCT image.