

## Employing Reasoning within the Phenoscape Knowledgebase

James P. Balhoff<sup>1,2</sup>, Wasila M. Dahdul<sup>3</sup>, Hilmar Lapp<sup>1</sup>, Peter E. Midford<sup>1</sup>,  
Todd J. Vision<sup>1,2</sup>, Monte Westerfield<sup>4</sup>, Paula M. Mabee<sup>3</sup>

<sup>1</sup>National Evolutionary Synthesis Center, Durham, NC, USA

<sup>2</sup>University of North Carolina, Chapel Hill, NC, USA

<sup>3</sup>University of South Dakota, Vermillion, SD, USA

<sup>4</sup>University of Oregon, Eugene, OR, USA

The Phenoscape project (<http://phenoscape.org>) links evolutionary phenotype descriptions to model organism phenotypes via ontological annotation of comparative data, with the ultimate goal of generating hypotheses about the genes involved in evolutionary phenotype transitions. To date, we have created ontological phenotype descriptions for over 11,000 evolutionary character states linked to 2500 ostariophysan fish taxa. We combined these annotations with mutant phenotype annotations derived from the Zebrafish Information Network (ZFIN), along with shared anatomy, quality, and taxonomy ontologies, into the Phenoscape Knowledgebase. The utility of ontological annotations is derived from the use of shared identifiers for concepts and the explicit computable semantics of those concepts provided by the ontologies. By building on a semantic data store and reasoning system (OBD, the Ontology-Based Database), the Phenoscape Knowledgebase allows users of its web interface to exploit the rich semantics of phenotypic annotations.

Because phenotypes are expressed as complex intersections of anatomy and quality (and often other) ontology terms, the Knowledgebase relies on automatic classification by the OBD reasoner to infer subsumption of phenotypes by other phenotypes as well as placement within the core ontology hierarchies. Class subsumption, transitive properties, and property chains are all employed in providing results within the

Knowledgebase web query interface. For example, users are able to specify a search for all phenotype annotations involving the "shape of any part of the head, within the taxon Cypriniformes". By default, queries within higher taxa (e.g. an order such as Cypriniformes) return annotations to any of their component subtaxa. However, via an option in the query interface, users can make use of additional semantics to automatically infer that phenotypes annotated to a higher taxon apply to all its members.

The OBD reasoner provides an interface for custom rules to be incorporated into the reasoning framework. Using this interface, the Phenoscape Knowledgebase allows automatic inference of absence phenotypes for structures which develop from (as stated in the anatomy ontology) structures which have been asserted as absent within a given taxon. Using the reasoning framework and rule interface allows domain-specific assumptions such as these to be stated explicitly and not buried in application code.

Phenoscape is continuing to address other reasoning challenges. We are developing a logical framework for homology assertions which will allow proposed homologies to be correctly propagated in the context of anatomy ontologies and phenotype annotations. Additionally, we are exploring methods for incorporating taxon-specific anatomical relations into our anatomy ontology which can facilitate queries within specific taxonomic contexts.