# COLLABORATIVE DEVELOPMENT OF A NEW VERTEBRATE TRAIT ONTOLOGY FOR COMPARISON OF GENOTYPE AND PHENOTYPE DATA ACROSS SPECIES


***Diane Munzenmaier[1], Susan Bello[2], Cynthia Smith[2], Carissa Park[3], Mary Shimoyama[1], Janan Eppig[2], and James Reecy[3]***

*[1]Rat Genome Database (RGD), Medical College of Wisconsin, Milwaukee, Wisconsin*
*[2]Mouse Genome Database (MGD), The Jackson Laboratory, Bar Harbor, Maine*
*[3]Animal QTL Database (Animal QTLdb), Iowa State University, Ames, Iowa*



## Abstract

With the increasing efforts to identify new QTL, describe strain characteristics, and use genetic data to make cross-species comparisons, the need to develop a standardized ontology to define and classify vertebrate traits has become urgent. To meet this need, RGD, MGD and Animal QTLdb have joined forces to build the Vertebrate Trait Ontology (VTO). The purpose of the VTO is to represent and compare traits (defined as "any measurable or observable characteristic" related to an organism, organ, tissue, or cell) between animal strains as well as across animal species. The VTO is designed to allow for consistent annotation by curators and ease of use. This new ontology was initially adapted from the Mammalian Phenotype (MP) Ontology which represents abnormal mammalian phenotypes, terms from the Animal Trait Ontology have been incorporated, and the overall structure has been modified to reflect the broader scope of this ontology. The VTO will eventually be integrated with ontologies for experimental conditions and clinical measurements so that data acquired using various techniques from many different laboratories as well as clinical data can be standardized, allowing more efficient access to and analysis of relevant genotype and phenotype data.



**REFERENCES**

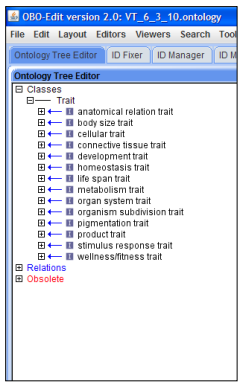
Hughes, LaRon M., J. Bao, Zhiliang Hu, Vasant G. Honavar, James M. Reecy (2008). Animal Trait Ontology (ATO): The importance and usefulness of a unified trait vocabulary for animal species. *Journal of Animal Science, 86:1485-1491.*

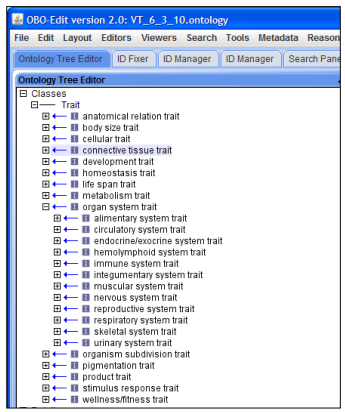
Shimoyama, Mary, Victoria Petri, Dean Pasko, Susan Bromberg, Wenhua Wu, Jiali Chen, Nataliya Nenasheva, Anne Kwitek, Simon Twigger, Howard Jacob (2005). Using multiple ontologies to integrate complex biological data, *Comparative and Functional Genomics, 6:373-378.*

Smith, Cynthia L., Janan T. Eppig. The mammalian phenotype ontology: enabling robust annotation and comparative analysis (2009). *Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 1:390-399.*


### What is the Vertebrate Trait Ontology and why do we need it?

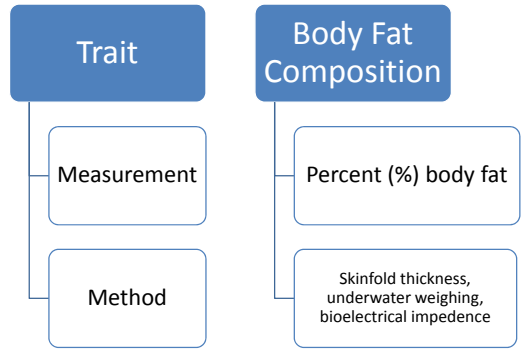
The use of quantitative trait loci in order to correlate of a property of an organism with a genetic sequence variation is becoming widespread. As the field develops, nomenclature is becoming more disease-focused and species-specific and the distinction between trait and measurement is becoming blurred. The Mammalian Phenotype ontology was originally designed for abnormal or disease phenotypes while the Animal Trait Ontology was specifically intended to describe livestock traits for agricultural use. Thus, the need exists for a standardized resource that will assist clinical and basic science researchers in the search for the genetic basis of biological variation. Our Vertebrate Trait Ontology (VTO) is an attempt is to generate a common vocabulary to better reflect trait variations within and between species in both health and disease.




### Normal variation vs. Disease

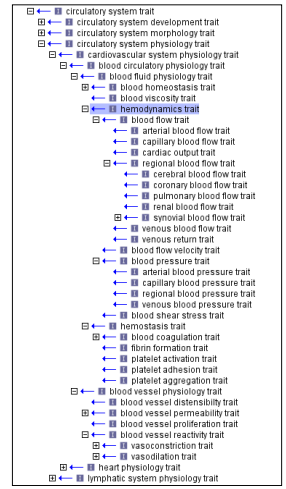
The Mammalian Phenotype (MP) ontology was developed with the goal of categorizing abnormal phenotypes resulting from gene knockout experiments. However, we know that normal, healthy organisms exhibit a large amount of morphological and physiological variation that is often due to genetic variation. Environmental factors, types of measurements, methods, and conditions also can play a role in expressed phenotypes which can complicate analysis. The VTO is an attempt to standardize the attributes that are assessed across species and strains so that valid comparisons can be made which might lead to more informative conclusions regarding the genetic basis of structure and function within and between organisms, strains, and species. This will allow more accurate assessment of the genetic basis for biological variation as well as facilitating the identification of optimal research strains and disease models.




### Trait vs. Measurement

An important distinction in the VTO is how a trait differs from a measurement of a trait. A trait is a (morphologic or physiologic) characteristic or property of the organism that the researcher is interested in. A measurement is a quantitation or assessment of a specific trait. A method is the manner in which the measurement data is acquired. For every trait, there can be numerous types of measurements that can be used to try to assess the trait. For every measurement, there can be various methods that are in standard use.

For example, body fat composition has been shown to be strongly correlated to disease which makes it a valuable trait to study. One common assessment or measurement of body fat composition is "percent body fat". This can be measured by skinfold thickness using a specialized caliper, underwater weighing, or bioelectrical impedence. Each method has its own caveats and built-in error but each is used to determine a measurement of the same trait. Body Mass Index (BMI) is another means of assessing body fat composition but that is another type of measurement (index vs % fat) using yet another method with its own caveats.




The VTO is divided into traits associated with development, morphology, and physiology for each major organ system as well as other major divisions. These general categories are reduced to more detailed descriptions as required to incorporate distinct traits associated with trait measurements reported in the literature. This logical structure will allow for simple and frequent updates and expansion of the ontology as necessary.

The VTO is a work in progress which is scheduled to be released later this year. Updates are expected to be done on a regular basis over the next 2-3 years.

### Conclusions and Future Directions

The VTO has been designed to serve as a benchmark vocabulary for the assessment of biological variation within and between all vertebrate strains and species. Use of this standardized format will help to reduce the inconsistencies currently found in the literature in regard to references to various traits and the measurements used to assess traits. It will also allow easier comparison of traits between species, including humans, because trait descriptions are standardized and not species-specific.

Additional ontology development is currently underway for clinical measurements, methods, and experimental conditions, which will be used in parallel with the VTO to further define and standardize how traits are being assessed and to enable more informative conclusions to be made from the data.


This project is supported by funds from the USDA, Cooperative State Research, Education, and Extension Service (CSREES) Award 2008-35205-18765.

The Rat Genome Database is funded by NIH 5R01HL064541.


United States Department of Agriculture

National Institute of Food and Agriculture