# Animal QTLdb Extension (V): Addition of New Data Types and Functions

Zhi-Liang Hu, Carissa Park, and James M. Reecy

Department of Animal Science and Center for Integrated Animal Genomics, Iowa State University, 2255 Kildee Hall, Ames, IA 50011

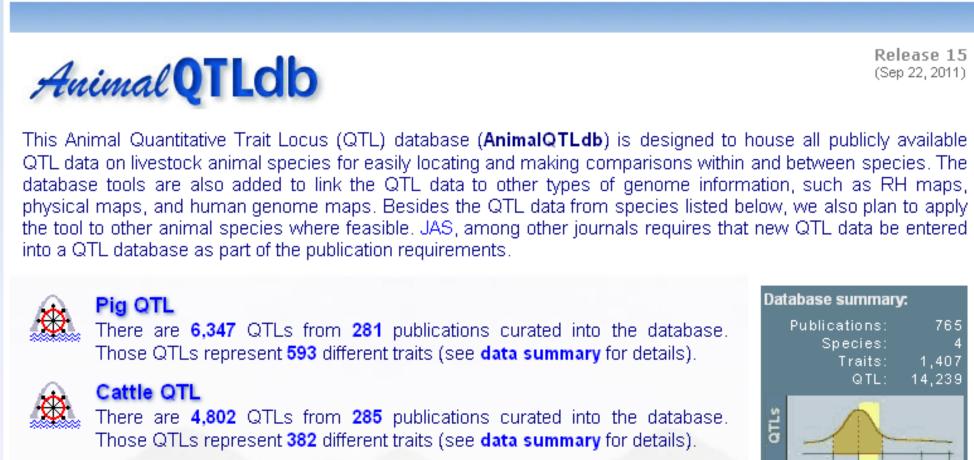
### Abstract

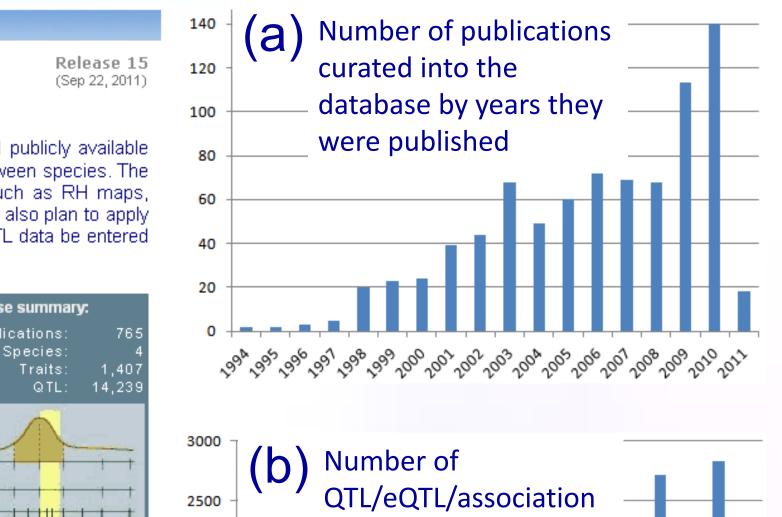
The Animal QTL Database (QTLdb) has been developed to provide a centralized resource tool to link traits to respective genome locations using curated data in the public domain. To date, it is housing more than 15,000 QTL in 5 animal species (cattle, chicken, pig, sheep and rainbow trout). The Animal QTLdb was built not only to house QTL data for within- and between-species comparisons, but also to facilitate comparative alignment with genome features such as genes, transcripts, RH markers, SNPs, etc. Rapid progress in genome-wide association studies (GWAS) and construction of the Vertebrate Trait Ontology have presented new opportunities for further development of the Animal QTLdb. We report here our recent efforts on the QTLdb developments which have made it possible to accommodate and combine the GWAS study results, and to link out from the QTLdb to the newly constructed Vertebrate Trait, Livestock Product Trait, Clinical Measurement, and livestock breed ontologies. We consider these developments important steps to standardize the ways in which various databases can directly "talk" to each other and exchange information. Our continued efforts to develop the QTLdb are to ensure it is up-todate in terms of both data contents and functions, and also to make it a central hub bridging between genotypes and phenotypes for all livestock species, which will further facilitate comparative studies across species.

# http://www.AnimalGenome.org/QTLdb

EOMICONINACIAN CONTROL

**Figure 1.** A snapshot of the Animal QTLdb homepage as well as charts showing the growth of the database in both increased number of curated publications (a), and the number of QTL/ eQTL/ association data added.





# Terminology

#### "QTL" data:

We generally refer to "QTL" data as data from both QTL and association studies, being it a genome/chromosome-wide scan, an interval mapping, or results from other types of statistical analysis.

#### The Vertebrate Trait Ontology (VT) is a controlled vocabulary for the description of traits (measurable or observable characteristics) pertaining to the morphology, physiology, or development of vertebrate organisms

#### **The Clinical Measurement Ontology** (CMO) is designed to be used to

# Introduction

The AnimalQTLdb was designed to house all publicly available QTL data on livestock animal species for easily locating data and making comparisons within and between species. The functions of the database were subsequently expanded by adding tools to link the QTL data to other types of structural genomic information, allowing comparative viewing of RH maps, physical maps, SNPs, microarray elements and human genome maps (Hu et al., 2007). The QTLdb was further developed to transfer QTL information from linkage map to its respective genome map and display it with Gbrowse (Hu et al., This research tool to link between phenotypes and 2009). genotypes has greatly facilitated the community research activity on positional mining of QTL information (Hu et al., 2010). The QTLdb has been cited in more than 150 journal papers during the past 7 years.

As more genome assemblies for livestock species become available, and high-density SNPs are used for phenotype and genotype association studies (e.g., genome-wide association studies, GWAS), the need for the ability to mine genome annotation information for traits of interest and compare data between different map types continues to grow. In addition, demand for more standardized trait terminologies is high to facilitate comparisons across studies. Here we present our work on developing the QTLdb towards a more useful tool for researchers.

There are 2,451 QTLs from 125 publications curated into the database Those QTLs represent 248 different traits (see data summary for details).

#### Sheep QTL

Ð

Chicken QTL

There are 639 QTLs from 74 publications curated into the database. Those QTLs represent 184 different traits (see data summary for most recent updates)

#### Rainbow Trout QTL

There are QTLs from publications curated into the database. Those QTLs represent different traits (see **data summary** for most recent updates)

Frequently asked questions: Questions and answers designed to serve as a User's Guide to QTLdb, addressing the usage, application, utilities, new developments and related issues.

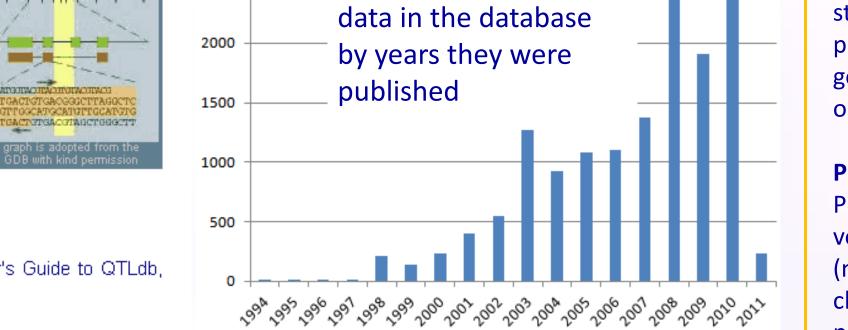
New Data: If you have new QTL data that has not been included in the current release, please either (1) register as a QTLdb curator to input it directly by yourself (see "Database Curators / Editors" below), or (2) inform our Bioinformatics Team so that we can include it for curation before the next release.

Database Curators / Editors: The QTLdb is open to public for data entry and update. One can also use the QTLdb as a research tool by keeping your own data in private mode.

**Data Downloads:** You can download the QTLdb data at each species front page, or at individual chromosome page. The QTL location data available for download include "cM" in tab delimited form and "bp" in GFF format. See FAQ #22 for details.

## **Results** (continued)

- 4. A number of new features are now programmed into the QTLdb tools to better facilitate the database usage (see Figure 3).
  - a) Options to view different data types, allowing "QTL", "eQTL", and "association" data to be displayed in various combinations.
  - b) Options to view "Linkage Map" or "Genome Map" or both.
  - c) Options to view different genome map builds if applicable.
  - d) Cytogenetic bands and genome coordinates are aligned side by side with linkage map.
  - e) New shapes/colors for different data types.



standardize morphological and physiological measurement records generated from clinical and model organism research and health programs.

**Product Trait Ontology** (PTO) The Product Trait Ontology is a controlled vocabulary for the description of traits (measurable or observable characteristics) pertaining to products produced by or obtained from the body of an agricultural animal or bird maintained for use and profit.

#### Figure 3.

A screenshot of an Animal QTLdb chromosomal map view, showing new functions added to the viewer: (a) Options to view different data types, allowing "QTL", "eQTL", and "association" data to be displayed in various combinations; options to view "linkage map" or "genome map" or both; and options to view different genome map builds if applicable. (b) Cytogenetic bands and genome coordinates aligned side by side with linkage map. (c) New shapes/colors for different data types. In addition, all QTL/traits found on a chromosome are now listed in a pull-down menu for users to easily browse and view them (d).

	CattleQTLdb			Browse Search	View Maps FAQ
$(a)_{\epsilon}$	Search:	Chromosome: 1 💌 Sort by: Trait	s 🛛 💙 Zoom: Small 👌	Markers: Filled	Y Font size: Tiny Y
a)<	🛛 Data: 🗹 QTL 🗹 eQTL 🗹 As	ssociation Map: 🗹 Link	age Map 🛛 🗹 Genome Map	Build: d	oSNP.Btau_4.2 💌
	A Snapshot of QT	L on Cattle Chrom	osome 1 🛛 🗖	lign QTL from:	(cM) to: (cM)
	( There are too many QTL to display;	Use the browse hint below the graph to	1 - T	Baylor Cattle SNPs alig	ned to bovine genome 🛛 👻
	QTL Mapper v.2.019	t.			
			8		
b)	91.1 -BM8139 -10 91.3 20 -DIK4957 -20	· 1 · · ·		,8 <u>,</u>	18
	92.1 10 - DVEPC63 -40	a 🕹 🔊	B	10 JOO A0	V D0PV DRIPL DRIPL
	92.5 92.5 93.1 BMS4001 −70		CHOPT CONCEPT	÷.	
	43.3 90	. <mark>.</mark> .		· · ·	
	93.5 - 100 - BMS1170 - 110	<sup>10</sup>	CHHO CHHO		l l l l l l l l l l l l l l l l l l l

# Results

Recent development of the Animal QTLdb has been mainly focused on inclusion of more data types to facilitate integrated genome analysis.

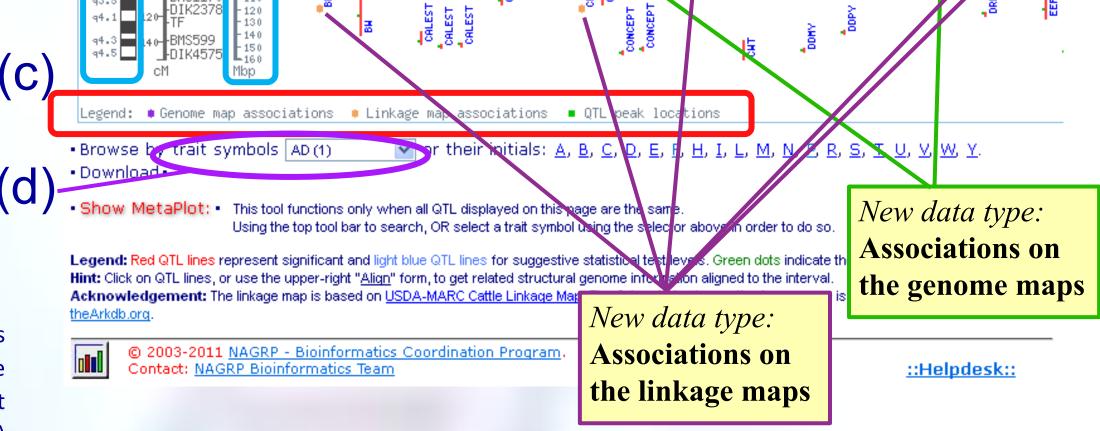
- 1. Continued curation of new data into the QTLdb. On average, we have been making 3 database releases every year in the past few years. As of *Release 16*, there have been **15,014** "QTL" data curated from 784 publications in 5 species, which represent **1,413** animal traits. In **Figure 1** is shown the growth trends of the database in terms of increased number of curated publications and the number of QTL/eQTL/association data added. Note that sheep and rainbow trout have become two new animal species served by the QTLdb.
- 2. A number of new data types and parameters are made available for data collection into the database. The new data analysis types include "QTL", "eQTL" and "association". The addition of "association" data type is significant, because in the near future GWAS data will likely be more dominant in the study of phenotype/genotype relationships. The newly added test models are "paternally imprinted", "maternally imprinted", "imprinted", "sex-specific", "epistatic" and "Mendelian" (Figure 2). The

- All QTL/traits found on a chromosome are now listed in a pulldown menu for users to easily browse and view them.
- g) The graph resolution of chromosomal views is improved to meet publication quality.

#### Figure 2.

A snapshot of an Animal QTLdb data details page, showing new parameters and/or features added to the database. The new parameters subject to data collection into the database include data analysis types (c), test models (c), and animal breeds (b). The Animal Trait Ontology (ATO) is now linked to Vertebrate Trait Ontology (VT), Product Trait Ontology (PT) and Clinical Measurement Ontology (CMO).

QTL	. Trait Information					
QTL Trait: Carcass weight			Vertebrate Trait Ontology: n/a			
QTL Symbol: CWT				Product Trait Ontology:         Dressed carcass weight           Clinical Measurement Ontology:         Carcass weight		
QTL	. Map Information			QTL Experiment in Brief		
Chromosome: 1			Animals: Twenty-nine generations of commercial American			
	QTL Peak Location:	148 (cM)	- Angus. - Breeds that this QTL is associated with: Angus			
	QTL Span:	139.3-154.6 (cM)				
ers	Upper, "Suggestive":	n/a		Design: Genome scan with 402 markers at 8 cM intervals or		
markers	Upper, "Significant":	BMS599		14 cattle production traits using 3391 animals from		
u Bu	Peak:	BMS4014		Ana Analysis Types:		
Flanking	Lower, "Significant":	#		"QTL"/"eQTL"/"association"		
Ē	Lower, "Suggestive":	n/a				
	Analysis type:	QTL		Software: GridQTL, LOKI		
C Model tested: Mendelian Test base. Genome-wise		Notes: Not a cross based segregation analysis Links: Edit				
	Threshold significance level:	Significant		Reference		
tics	LOD Score:	n/a		Authors: McClure MC, Morsci NS, Schnabel RD, Kim JW,		
Statistics	Least Square Mean:	n/a		Yao P, Rolf MM, McKay SD, Gregg SJ, Chapple RH, Northcutt SL, Taylor JE		
ti	P-values:	4.69	Te	est models:		
E L	F-Statistics:	n/a		Paternally imprinted"/"Maternally		
ļ	Variance:	liva 🛉				
	Bayes value:			nprinted"/"Imprinted"/"Sex-		
	Likelihood Ratio:	n/a	sp	pecific"/"Epistatic"/"Mendelian"		
	Dominance effect:	0.0132		Journal, Anim Genet.		
Additive effect: n/a		n/a		Issues: 41 (6): 597-607		
Candidate Gene: n/a			Links: No PubMed entry   Abstract   Paper   Edit			
	Links: NCBI GeneDB	Edit   View		Additional Information		
				Comments: SNP scan		



## Discussion

The improvement to Animal QTLdb has been a continual process since its inception in 2004. Given the rapid progress in animal genome research, it has been our long term goal to keep up with the demands of the community. Besides the recent accomplishments presented here, we also have a list of improvements we plan to implement in the Animal QTLdb in the near future.

In summary, our work extending the utility of the QTLdb is a significant step forward towards making the QTLdb a more useful research tool for the community (Hu and Reecy, 2007b).

## References

Hu, Zhi-Liang, Svetlana Dracheva, Wonhee Jang, Donna Maglott, John Bastiaansen, Max F. Rothschild and James M. Reecy (2005). A QTL resource and comparison tool for pigs: PigQTLDB. Mammalian Genome

addition of breed information will be useful for future data analysis to trace trait variations to animal ancestors by breeds. The addition of new data types and parameters is useful for future meta-analysis of stored data.

3. The Animal Trait Ontology (ATO) was developed as part of Animal QTLdb to help with managing and organizing trait information (Hu et al., 2005, 2007). Its data was incorporated into the Vertebrate Trait Ontology (VT), Product Trait Ontology (PT) and Clinical Measurement Ontology (CMO) by a consortium between the Animal QTLdb, the Rat Genome Database, and Mouse Genome Informatics. The mapping between the ATO and VT/PT/CMO (Figure 2) is one of the steps to formalize the trait ontology for more consistent data analysis results. The mapped VT/PT/CMO terms are hyper-linked to the respective hierarchy tree display using AmiGO (URL: http://amigo.geneontology.org).

# Acknowledgements

#### 16(10):792-800.

Hu, Zhi-Liang, Eric Ryan Fritz and James M. Reecy (2007a). AnimalQTLdb: a livestock QTL database tool set for positional QTL information mining and beyond. Nucleic Acids Research, 35(Database issue):D604-D609.

Hu, Zhi-Liang and James M. Reecy (2007b). Animal QTLdb: beyond a repository - a public platform for QTL comparisons and integration with diverse types of structural genomic information. Mammalian Genome, 18: 1-4.

Hu, Zhi-Liang, Cari Park, Eric Fritz, Mindy Dwindell, Mary Shimoyama, James M. Reecy (2009). Expanding the Utility of Animal QTLdb. Plant & Animal Genomes XVII Conference, January 10-14, 2008, Town & Country Convention Center, San Diego, CA.

Hu, Zhi-Liang, Carissa A. Park, Eric R. Fritz and James M. Reecy (2010). QTLdb: A Comprehensive Database Tool Building Bridges between Genotypes and Phenotypes. Invited Lecture with full paper published electronically on The 9th World Congress on Genetics Applied to Livestock Production. Leipzig, Germany August 1-6, 2010.

This research is supported by the USDA-NRI Grant number 2007-04187 and the USDA NAGRP Bioinformatics Coordination Project. We wish to thank James Koltes, Xiao-lin Wu and Eric Fritz for their useful discussions.