### Phenotype annotation

Jim Balhoff NESCent Informatics, Phenoscape

#### Annotation

- \* "Applying" ontologies to data
- \* Tagging
- \* Modeling Knowledge Representation

#### What is annotated

- \* Gene function/cellular location/process
- \* Gene expression patterns (anatomy)
- \* Phenotype (mutant, comparative, descriptive)

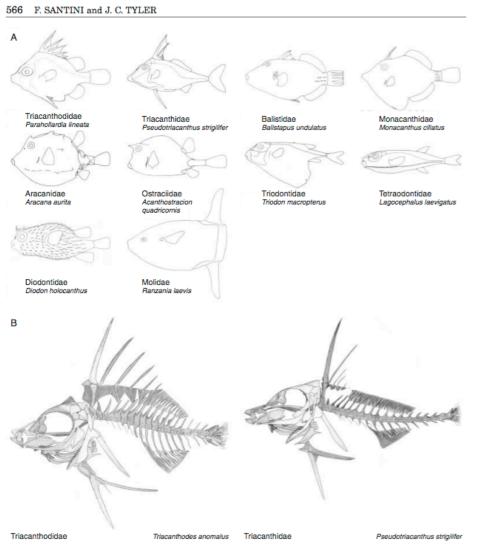
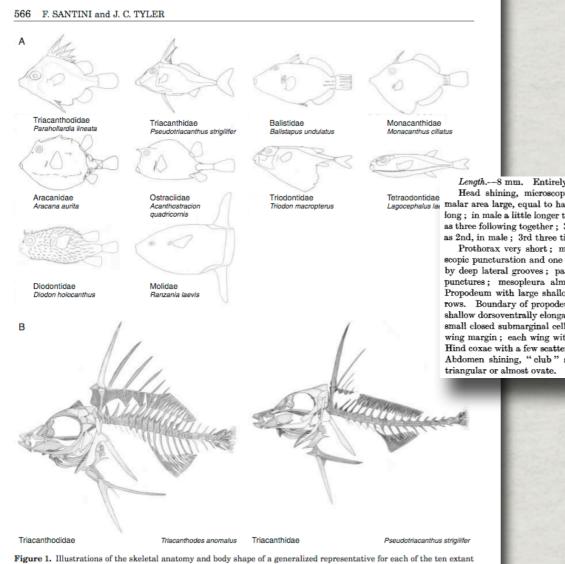


Figure 1. Illustrations of the skeletal anatomy and body shape of a generalized representative for each of the ten extant tetraodontiform families. Lateral views of (A) ten species, one from each extant family, and (B-F) of the skeletons of a representative species for each family, as follows: (B) triacanthodid Triacanthodes anomalus and triacanthid Pseudotriacanthus strigilifer; (C) balistid Balistapus undulatus and monacanthid Monacanthus ciliatus; (D) aracanid Kentrocapros aculeatus and ostraciid Acanthostracion quadricornis; (E) triodontid Triodon macropterus and tetraodontid Lagocephalus laevigatus; (F) diodontid Diodon holocanthus and molid Ranzania laevis.

© 2003 The Linnean Society of London, Zoological Journal of the Linnean Society, 2003, 139, 565-617



Length .-- 8 mm. Entirely black.

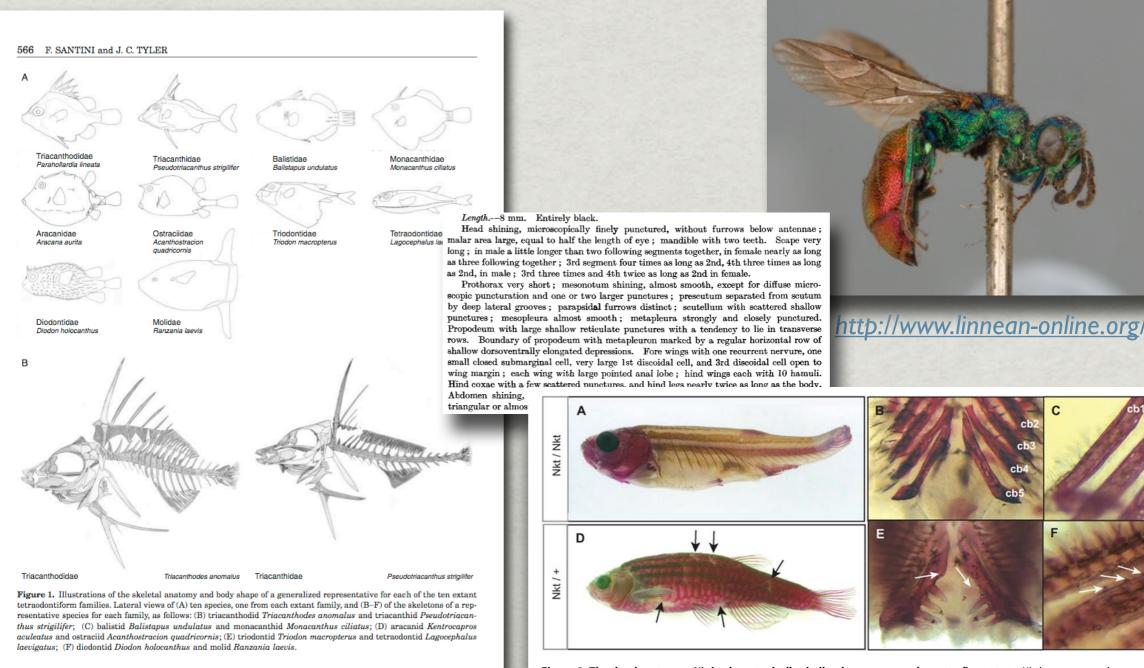
Head shining, microscopically finely punctured, without furrows below antennae; Lagocephalus lat malar area large, equal to half the length of eye; mandible with two teeth. Scape very  $\log$ ; in male a little longer than two following segments together, in female nearly as  $\log$ as three following together; 3rd segment four times as long as 2nd, 4th three times as long as 2nd, in male; 3rd three times and 4th twice as long as 2nd in female.

Prothorax very short; mesonotum shining, almost smooth, except for diffuse microscopic puncturation and one or two larger punctures; prescutum separated from scutum by deep lateral grooves; parapsidal furrows distinct; scutellum with scattered shallow punctures; mesopleura almost smooth; metapleura strongly and closely punctured. Propodeum with large shallow reticulate punctures with a tendency to lie in transverse rows. Boundary of propodeum with metapleuron marked by a regular horizontal row of shallow dorsoventrally elongated depressions. Fore wings with one recurrent nervure, one small closed submarginal cell, very large 1st discoidal cell, and 3rd discoidal cell open to wing margin; each wing with large pointed anal lobe; hind wings each with 10 hamuli. Hind coxae with a few scattered punctures, and hind legs nearly twice as long as the body. Abdomen shining, "club" sometimes reddish ventrally, strongly compressed laterally, triangular or almost ovate. Ovipositor not exserted.

http://www.linnean-online.org/

tetraodontiform families. Lateral views of (A) ten species, one from each extant family, and (B-F) of the skeletons of a representative species for each family, as follows: (B) triacanthodid Triacanthodes anomalus and triacanthid Pseudotriacanthus strigilifer; (C) balistid Balistapus undulatus and monacanthid Monacanthus ciliatus; (D) aracanid Kentrocapros  $aculeatus \ {\rm and} \ {\rm ostraciid} \ Acanthostracion \ quadricornis; (E) \ {\rm triodontid} \ Triodon \ macropterus \ {\rm and} \ {\rm tetraodontid} \ Lagocephalus$ laevigatus; (F) diodontid Diodon holocanthus and molid Ranzania laevis.

© 2003 The Linnean Society of London, Zoological Journal of the Linnean Society, 2003, 139, 565-617



© 2003 The Linnean Society of London, Zoological Journal of the Linnean Society, 2003, 139, 565-617

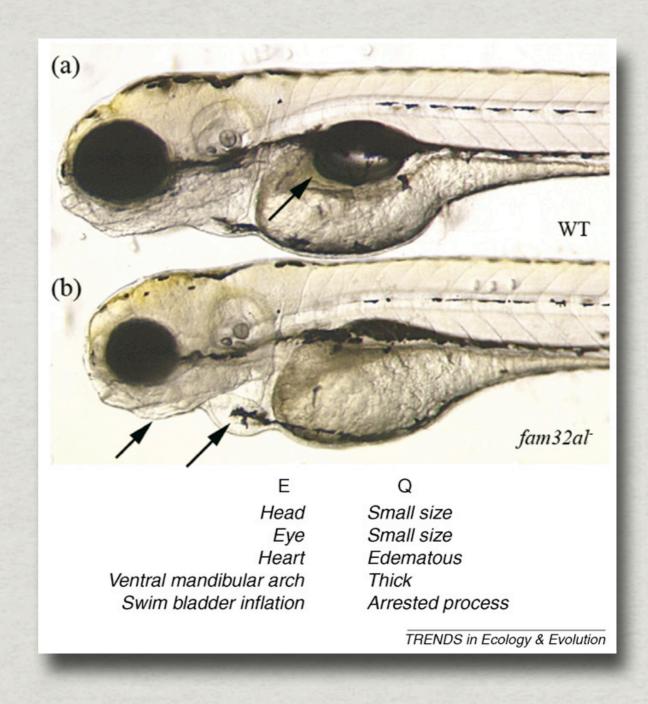
**Figure 2. The dominant gene** *Nkt* **is phenotypically similar, however complements** *fls* **mutants.** *Nkt* homozygotes show complete loss of scales, teeth and gill rakers resembling the *fls* phenotype (A–C). Heterozygous *Nkt* zebrafish show an intermediate phenotype of scale loss and patterning defect (arrows) while no effect on fin development is seen (D). Heterozygous *Nkt* also show a dominant effect on the number of teeth (arrows, E) and gill rakers (F), showing deficiencies along the posterior branchial arches and formation of rudimentary rakers along ceratobranchial 1 and 2 (arrows, F). *Cb1-5*, ceratobranchial bones. doi:10.1371/journal.pgen.1000206.q002

#### Phenotype annotation

- \* Ontology of phenotypes?
  - \* Can be done, but could become unwieldy
- \* Compositional approach: Entity-Quality

## Entity-Quality model

- \* Organism-specific anatomy entity terms are associated with more general phenotypic quality terms
- \* Compositional approach allowing any number of features to be described for a given entity



Brachyplatystoma capapretum: a New Species of Goliath Catfish from the Amazon Basin, with a Reclassification of

Allied Catfishes (Siluriformes: Pimelodidae) from Lundberg & Akama 2005

Table 1. Character State Matrix Used for Phylogenetic Analysis of the Placement of *B. capapretum* within Pimelodidae and *Brachyplatystoma*. Character states described in Appendix 1 and text.

	12345	67890	11111 12345	11112 67890	22222 12345	22223 67890	33333 12345
Steindachnerdion	11110	00000	00000	21000	00000	00000	00011
Phractocephalus-Leiarius group	11110	00000	00000	00000	00000	00000	00001
Pimelodus group	11111	11111	00000	00000	00000	00000	00000
Calophysus group	11111	11111	00000	00000	01110	00000	01201
Zungaro	11111	10000	00000	01001	20000	00000	00010
Sorubim group	11111	10000	00000	20001	20000	00000	00011
Platynematichthys	11111	10000	11000	00000	00000	00000	00000
Brachyplatystoma vaillantii	11111	10000	11111	11000	00000	00000	00000
B. tigrinum	11111	10000	11121	00111	11000	00000	10010
B. platynemum	11111	10000	11120	11111	11110	00000	11100
B. filamentosum	11111	10000	11111	11111	21101	11111	01101
B. capapretum	11111	10000	11111	11111	21101	11111	01101
B. rousseauxii	11111	10000	11111	11111	21101	11100	01101
Heptapteridae	00000	00000	00000	00001	00000	00000	00001
Pseudopimelodidae	00000	00000	00000	00000	01000	00000	10010
Bagridae	00100	00000	00000	01000	00000	00000	00000
Ictaluridae	00000	00000	00000	00001	00000	00000	00000

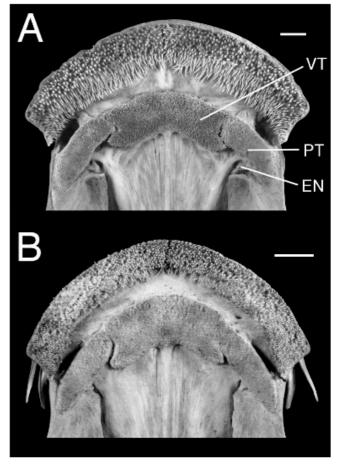
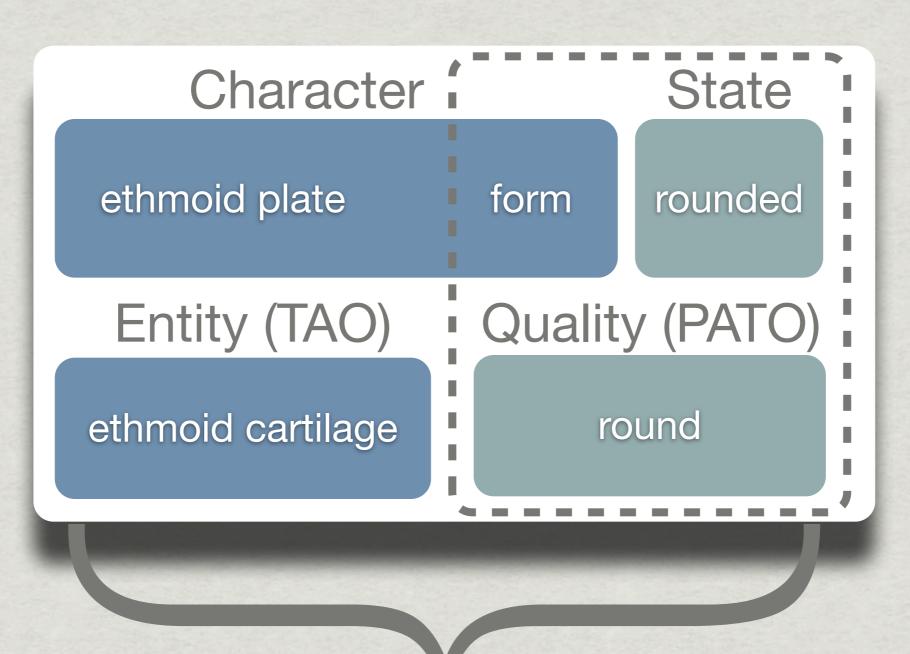


Fig. 7. Dentition of premaxilla and palate in ventral view of (A) *Brachyplatystoma filamentosum*, DU F1052; (B) *Brachyplatystoma capapretum*, MZUSP 53262. Scale bars = 1 cm. PT = pterygoid tooth plate, EN = endopterygoid, VT = vomerine teeth.

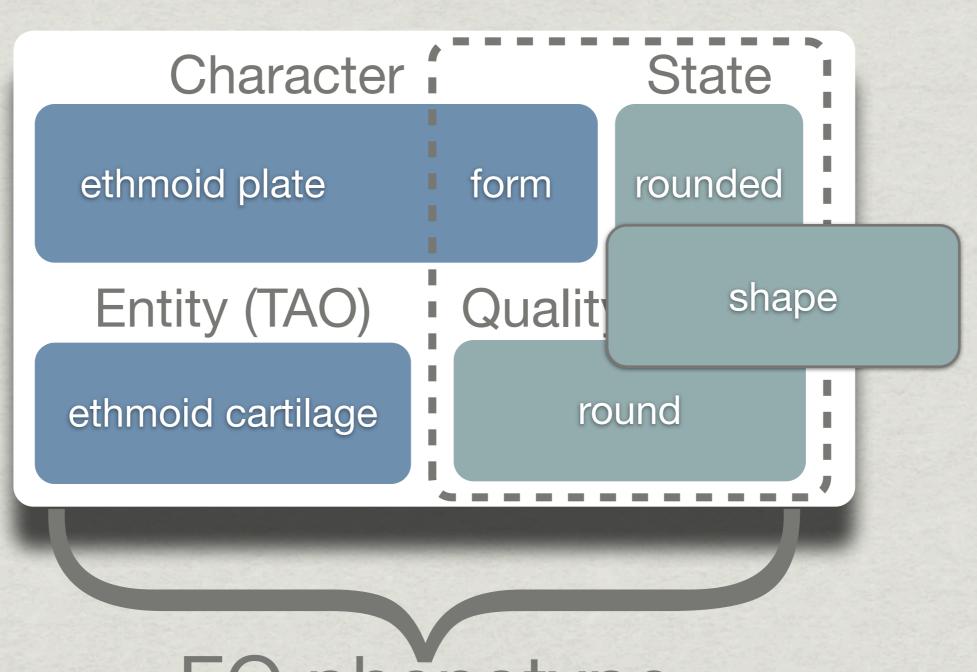
7. Ethmoid plate form: rounded [0]; quadrangular [1] (Lundberg et al., 1991b). Within Pimelodidae a uniquely derived and unreversed synapomorphy of *Pimelodus* group plus *Calophysus* group.

#### Entity-Quality Model for Evolutionary Phenotypes



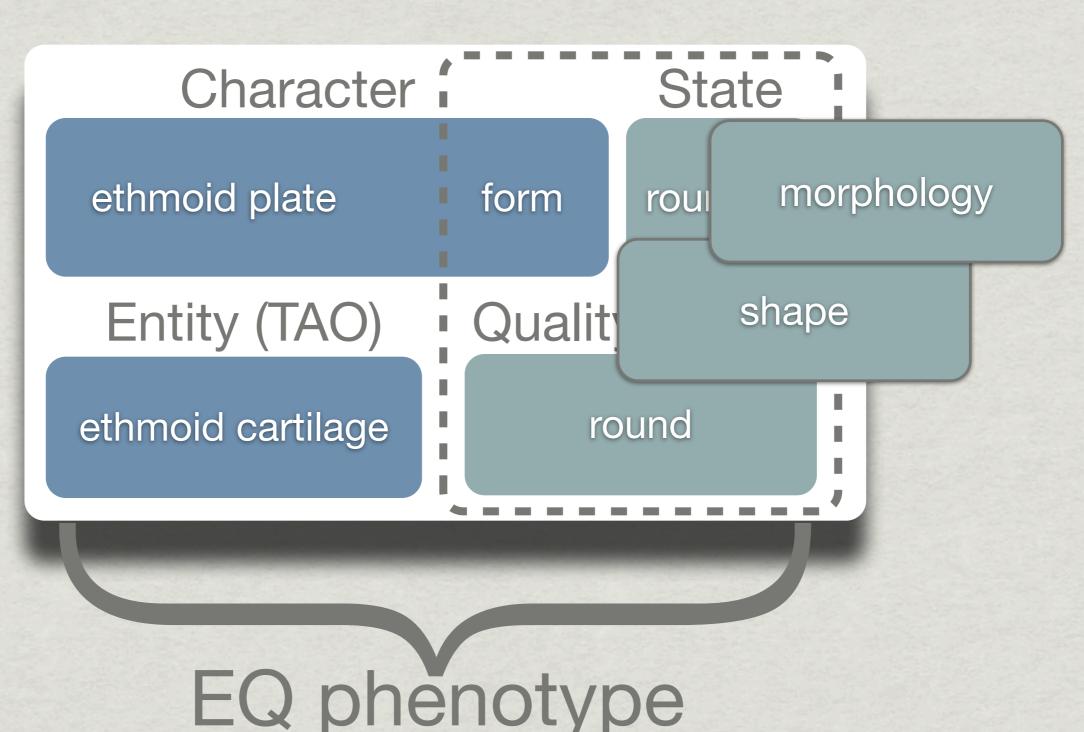
EQ phenotype

#### Entity-Quality Model for Evolutionary Phenotypes



EQ phenotype

#### Entity-Quality Model for Evolutionary Phenotypes



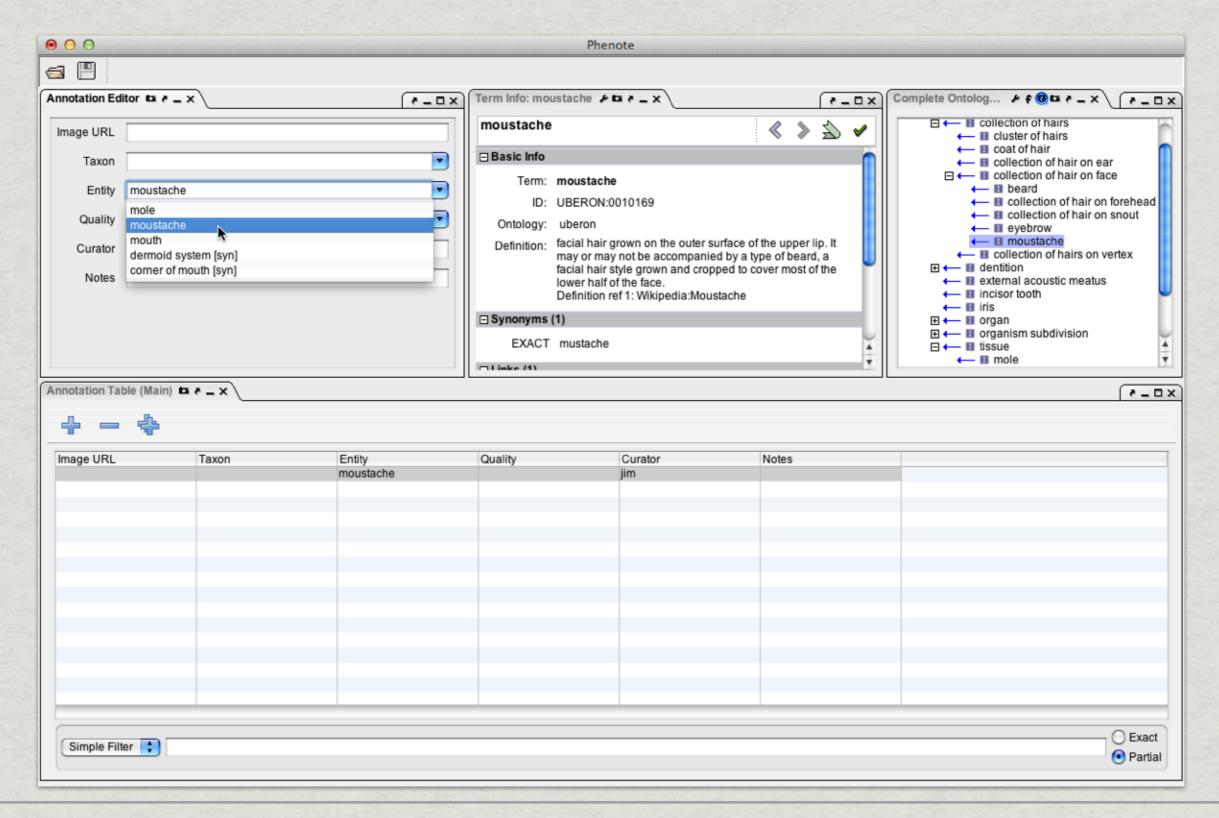
## PATO gotchas

- \* Deviation from "normal" ("increased size")
- \* Relational/dependent qualities ("sensitivity toward")

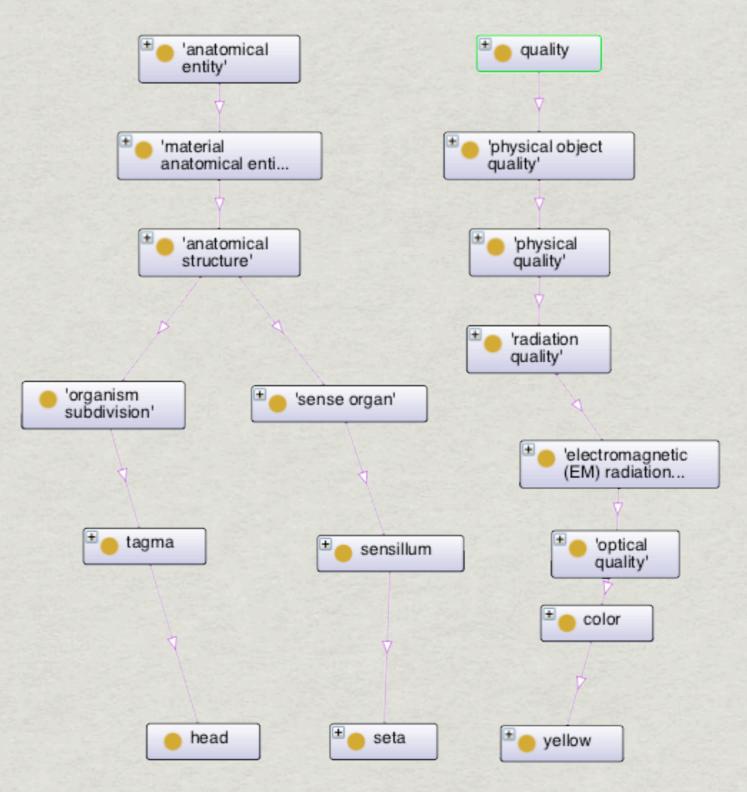
## Creating annotations

- \* Phenote generic, configurable table editor
  - \* loads ontologies, provides term autocomplete
- \* Phenex specialized for character matrix annotation
- \* Custom web interfaces
- \* Protégé OWL editor

#### Phenote



- \* Tagging/keyword provide search for data using ontology hierarchy
  - \* "Yellow bristles apparent on head"
  - \* tags: PATO:yellow, HAO:seta, HAO:head



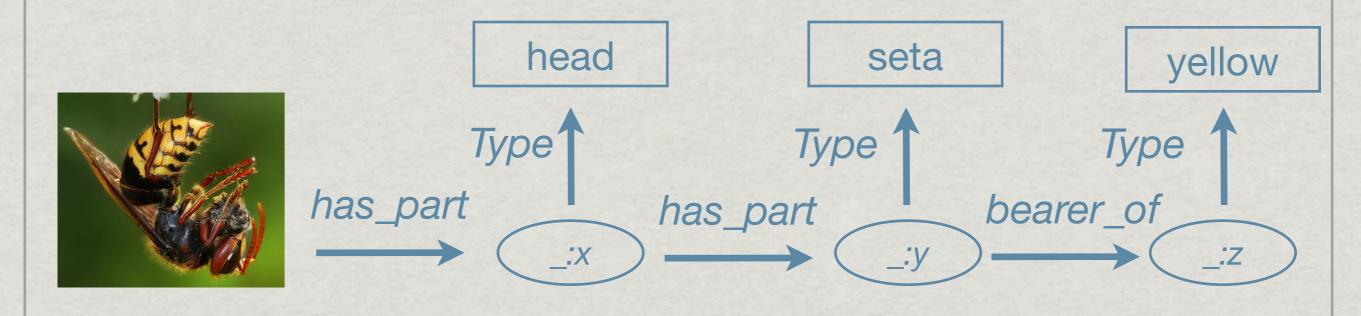
\* Can expand keyword query using hierarchy-aware middleware (e.g. -ontop- OBDA)

"Yellow bristles apparent on head"

- \* E and Q are not always sufficient
- \* Post-composition: E = "bristle which is part of head"
- \* L locator ("head")
- \* RE: Related/Dependent Entity needed for some PATO qualities ("sensitivity toward")

- \* Knowledge representation
  - \* Construct semantic model using EQ components
  - \* "Extend" ontology
  - \* Describe class of organisms or structures using a logical definition

- \* "Yellow bristles apparent on head"
- \* OWL class expression
  - \*\* has\_part some (head and has\_part some (seta
    and bearer\_of some yellow))



#### Limitations

- \* Difficult to adequately represent certain kinds of statements
  - \* "antennae absent" (works but complicates reasoning)
  - \* "antenna is longer than eye"
  - \* "increased count of antennae"
- \* Reasoning software does not scale well
- \* But, keep "tagging" utility in mind

#### Available annotations

- **\*** GO annotations
- \* Model organism phenotypes: ZFIN, MGI
- \* Virtual Fly Brain
- \* Phenoscape Knowledgebase
- \* Raw data available from most (all?)

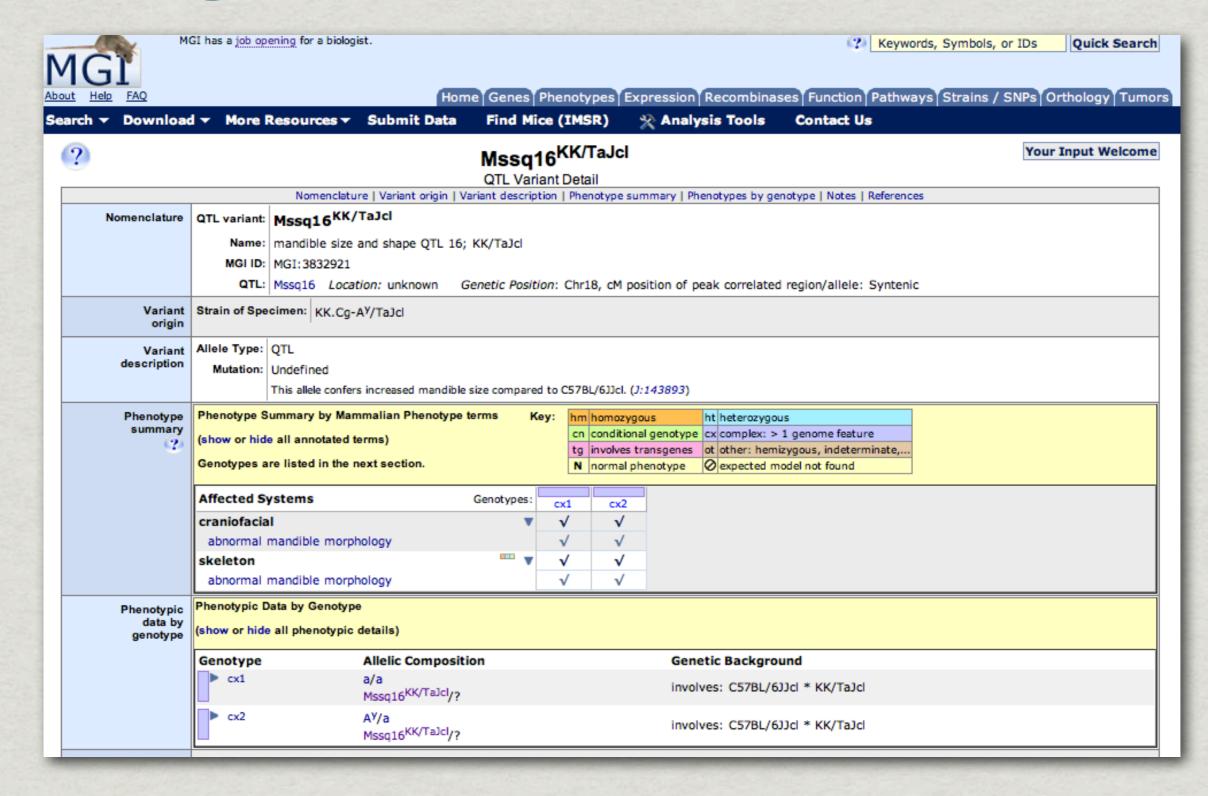
# Gene Ontology

PAA catabolism -  PAA pardinging  attribution  PAA pardinging  attribution  PAA page days  attribution  attribution  attribution  attribution  attribution  attribution  attribution  attribution  attribution  attribution	PA AN AND THE GE	ne Ont	ology	/				Ar	miGO
		Search	Browse	BLAST	Homolog /	Annotations	Tools & Resources	Help	
	Sea	arch GO			• terms	genes	or proteins exact mate	th Submit	
	mRNA transcription  Term associations ♣ Term information ➡ Term lineage ➡ External references ➡								
Ge	ene Product Associ	iations to	mRNA	transcr	iption;	GO:000	9299 and childre	า	
Dov	wnload all association infor	mation in:	gene ass	ociation for	mat 🗋 RDF	-XML			
1 2	Filter by Gene Product  Gene Product Type Data source Species  All  Complex  gene  gene  gene  gene  gene  gene  product  ASAP  Aspergillus fumig  Aspergillus niger   Tilter by Association  Evidence Code  All  IBA  IKR  IRD   Remove all filters  Remove all filters  Tilter by Association  Evidence Code  All  IRR  IRR  Remove all filters  Remove all filters						irs		
	Symbol, full n	ame		Informati	ion	Qualifier	Evidence	Reference	Assigned by
	cdk-7	10 associatio	-	<b>gene</b> from <i>Caenorhabditis</i> elegans			IDA	PMID:11960010	Wormbase
	ceh-37		ns <b>gene</b> f	gene from Caenorhabditis elegans			ISS	PMID:12711598	Wormbase
	Ereg epiregulin	39 associatio	<del>-</del>	protein from Mus musculus		UniProtKB:014944		MGI:MGI:4834177	MGI
	Ereg epiregulin			gene from Rattus norveaicus			ISS With UniProtKB:014944	RGD:1600115	UniProtKB (via RGD)

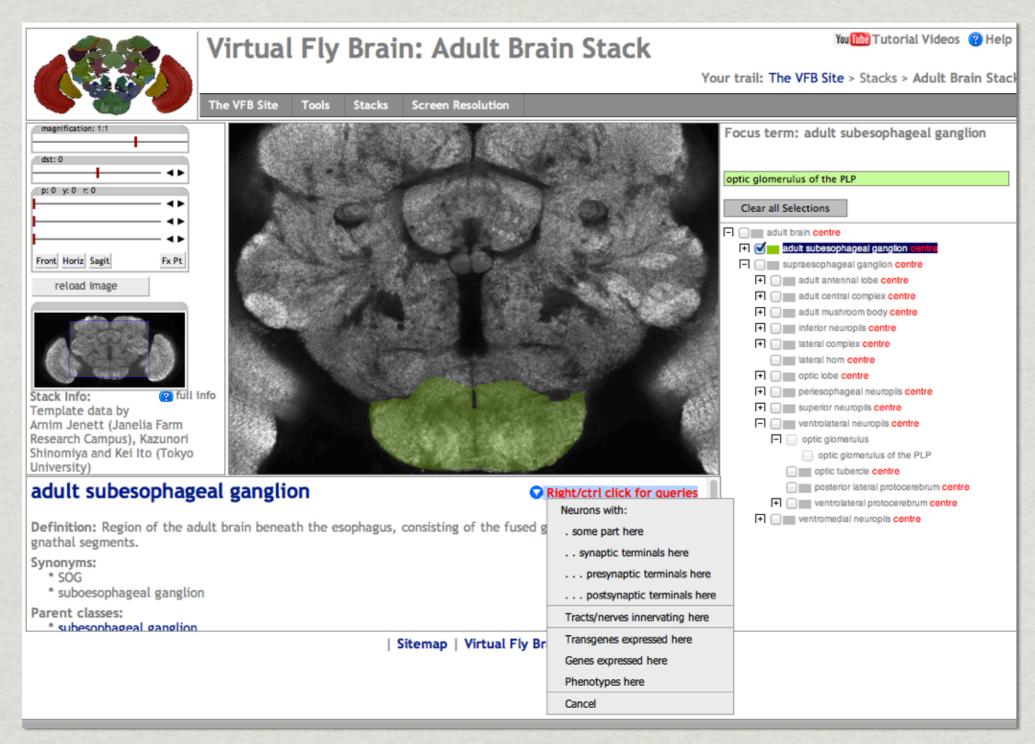


A THE	<b>.</b> T									Site Search:	
SZFI	N $\blacksquare$	Resear	ch	General	Information	ZIRC					
Home Genes / Marke	ers / Clones		GBrowse			Mutants / Morphants	/Tg An	atomy	Publications	Maps	
					ZFIN ID: ZD	B-GENO-080925-13					
											Your Input Welcome
	brpf1t20		02								
Background: Affected Gene:	Unspecifie brpf1	ed									
	No data a	vailable									
OFNOTVDE COM	DOCITION.										
GENOTYPE COMP		Zygosit	v		Parental G	Conotyne	Lab of	Origi	n	Construct	
t20002		homozy			♀+/- ♂+/-	enotype			nard Lab	Construct	
120002		Homozy	gous		+ 11-011-		rtussic	ATT-V OII	iaid Lab		
GENE EXPRESSION		1002/+2000	12								
Gene expression i			12			<b>6</b> ""					
Expressed Gene	Struc			bomere 5 □		Conditions standard or contr	ral	Figur		ave et el 2000	
egr2b hoxa2b						standard or contr		-		aue <i>et al.</i> , 2008 Laue <i>et al.</i> , 2008	1
TIOXEZE				! □ (all 6) ▶	car arcii 5-7	Staridard or Conti	01	o ngu	ies iii iioiii	Laue et al., 2000	,
			ch 2 □, pl 2 □ (all 6)	haryngeal a	rch 3-7 □,	chemical		Fig. 4	© from Lau	ue et al., 2008	
hoxb1a	rhomi	bomere 4				standard or control 3 figures 1 from		Laue et al., 2008	3		
		bomere 4				chemical		Fig. S	66 @ from L	aue et al., 2008	
Show all 7 express	ed genes										
PHENOTYPE 0											
Phenotype in brpf	f1t20002/t20	0002									
Phenotype						Conditions		Figur	es		
basihyal bone apla	stic, abnor	rmal 🗆				standard or contr	rol	Fig. S 2008	81 6 a , Fig.	S6 m2a, Fig. T1	from Laue et al.,
basihyal bone apla	stic, abnor	rmal 🗆				MO3-kat6a		Fig. S	66 🖾 from L	aue et al., 2008	
basihyal bone decr		-				standard or contr	rol	Fig. T	1 from Laue	e et al., 2008	
(normal or recovere						chemical				aue et al., 2008	
basihyal cartilage a						standard or contr	rol	Fig. 1	from Lau	ue <i>et al.</i> , 2008	
Show all 30 phenot	type stater	nents									

#### MGI



# Virtual Fly Brain



## Phenoscape KB



Site search:

Enter entity terms (e.g. basihyal bone), phenotypic qualities (e.g. shape, size), taxonomic names (e.g. lctaluridae), gene names or symbols (e.g. cadherin 6, cdh6), or publications.

Feedback

Home

Query taxa

Workspace

About

Acknowledgments

Query for:	
Taxon is:	Add
and	
Phenotype is O any O all:  fin • shape [broaden/refine]	Add
and	
Publication is ⊙ any ○ all : Any	Add

« Previous	Res	sults 1 - 20 of 56	Next »

<b>0 0</b>	Order ↓	Family	Taxon
	Siluriformes	Heptapteridae	Brachyrhamdia heteropleura
	Siluriformes	Heptapteridae	Genus 6 sp. (Bockmann 1998)
	Siluriformes	Heptapteridae	Phenacorhamdia unifasciata
	Siluriformes	Heptapteridae	Nemuroglanis pauciradiatus
	Siluriformes	Heptapteridae	Cetopsorhamdia molinae
	Siluriformes	Heptapteridae	Pimelodella sp. B (Bockmann 1998)
	Siluriformes	Heptapteridae	Chasmocranus chimantanus
	Siluriformes	Heptapteridae	Brachyrhamdia imitator
	Siluriformes	Heptapteridae	Mastiglanis asopos
	Siluriformes	Heptapteridae	Brachyrhamdia marthae
	Siluriformes	Heptapteridae	Brachyrhamdia meesi
	Siluriformes	Heptapteridae	Imparfinis guttatus
	Siluriformes	Heptapteridae	Horiomyzon retropinnatus
	Siluriformes	Heptapteridae	Phenacorhamdia sp. C (Bockmann 1998)
	Ciluriformoo	Hontontoridoo	Rimoladalla on A (Rookmann 1008)