Phenotype annotation

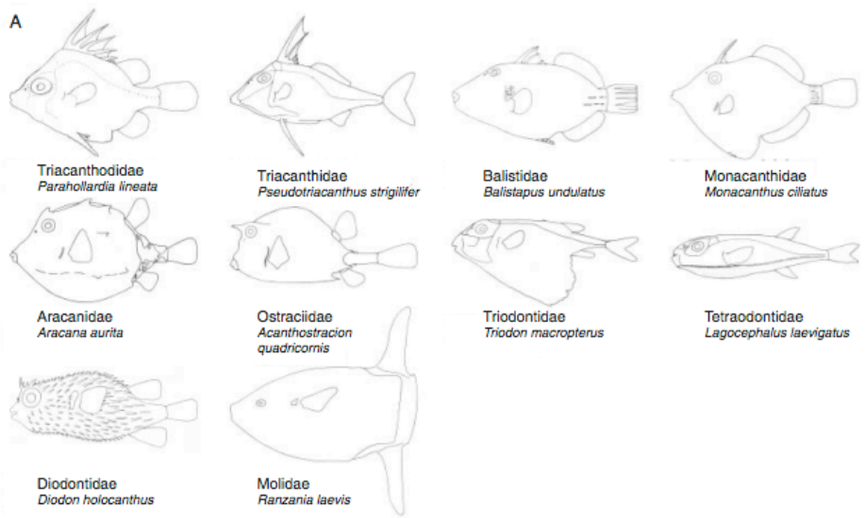
Jim Balhoff
NESCent Informatics, Phenoscape

Annotation

- “Applying” ontologies to data
- Tagging
- Modeling — Knowledge Representation

Phenotypic data

Phenotypic data



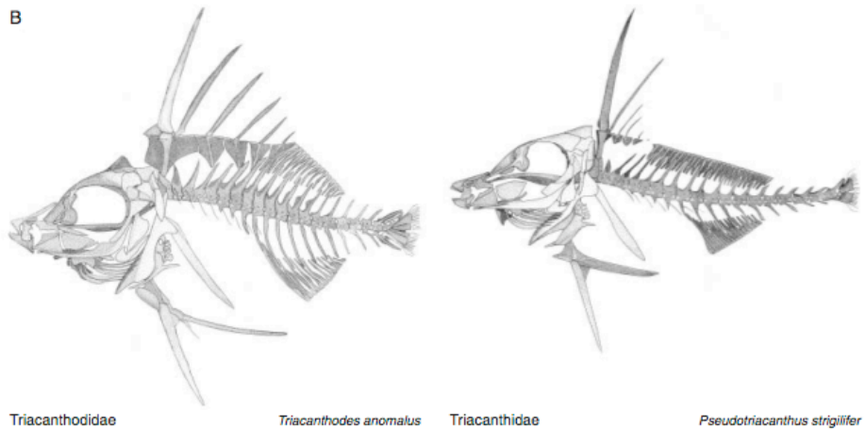
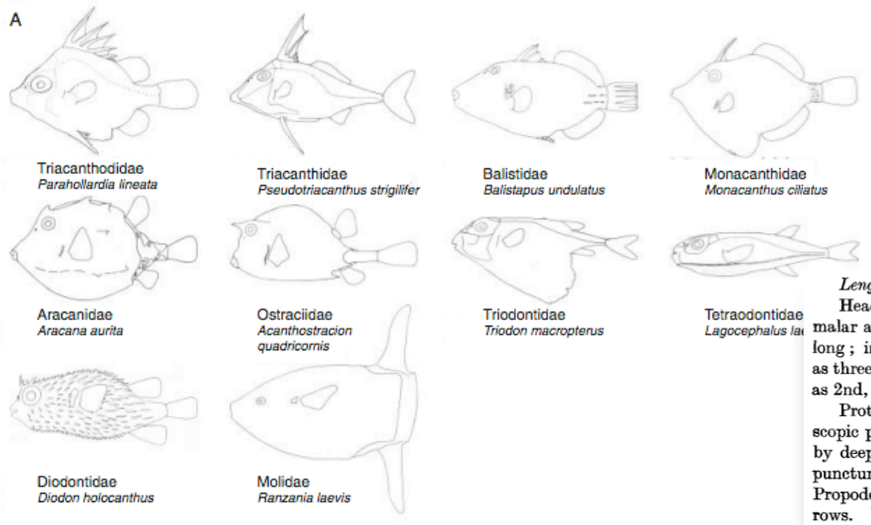


Figure 1. Illustrations of the skeletal anatomy and body shape of a generalized representative for each of the ten extant tetraodontiform families. Lateral views of (A) ten species, one from each extant family, and (B–F) of the skeletons of a representative species for each family, as follows: (B) triacanthodid *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

Phenotypic data

A



B

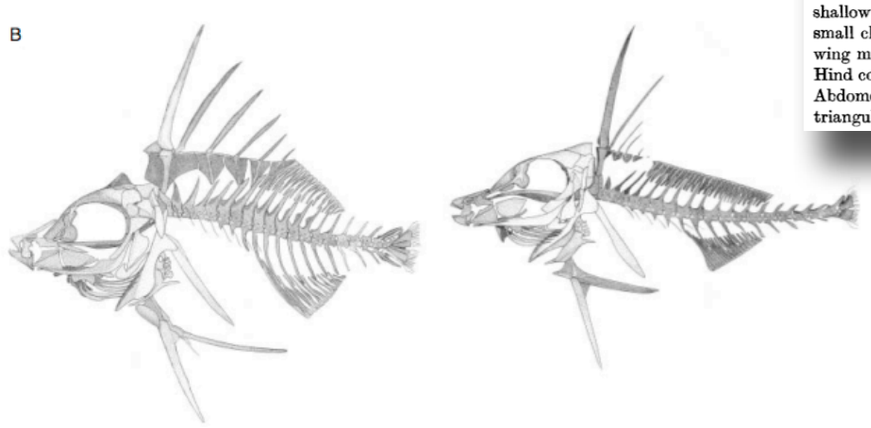


Figure 1. Illustrations of the skeletal anatomy and body shape of a generalized representative for each of the ten extant tetraodontiform families. Lateral views of (A) ten species, one from each extant family, and (B–F) of the skeletons of a representative species for each family, as follows: (B) triacanthodid *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

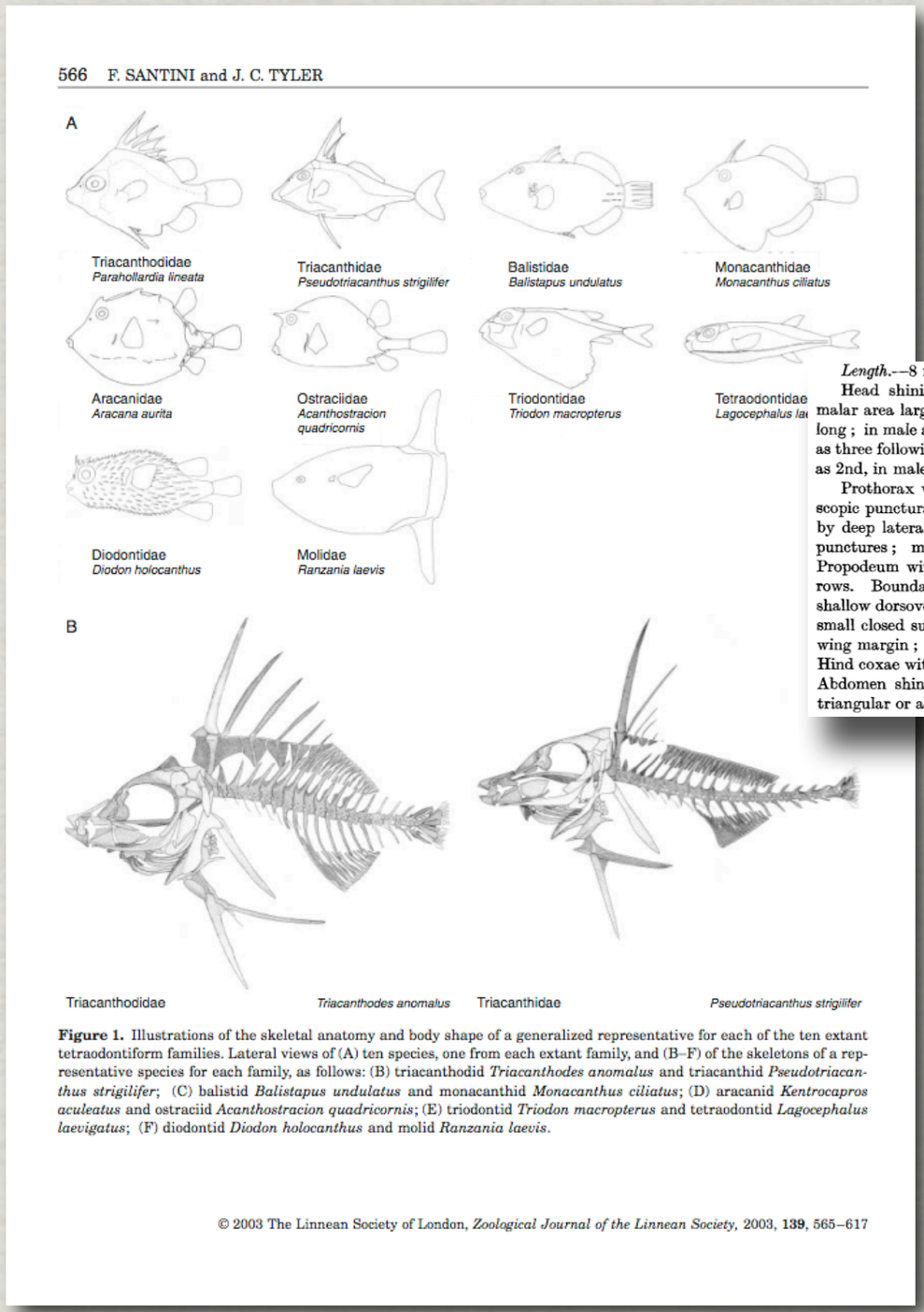
Length.—8 mm. Entirely black.

Head shining, microscopically finely punctured, without furrows below antennae; malar area large, equal to half the length of eye; mandible with two teeth. Scape very long; in male a little longer than two following segments together, in female nearly as long 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/

Phenotypic data



566 F. SANTINI and J. C. TYLER

A

Triacanthodidae *Parahollardia lineata*

Triacanthidae *Pseudotriacanthus strigilifer*

Balistidae *Balistapus undulatus*

Monacanthidae *Monacanthus ciliatus*

Aracanidae *Aracana aurita*

Ostraciidae *Acanthostracion quadricornis*

Triodontidae *Triodon macropterus*

Tetraodontidae *Lagocephalus la*

Diodontidae *Diodon holocanthus*

Molidae *Ranzania laevis*

B

Triacanthodidae *Triacanthodes anomalus*

Triacanthidae *Pseudotriacanthus strigilifer*

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; malar area large, equal to half the length of eye; mandible with two teeth. Scape very long; in male a little longer than two following segments together, in female nearly as long 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, triangular or almos

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

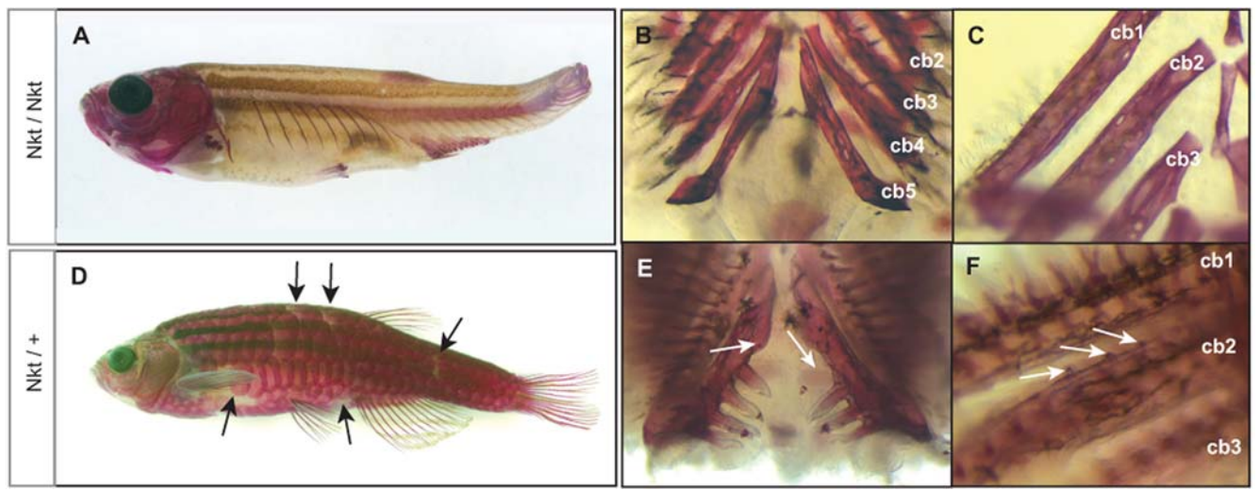


Figure 2. The dominant gene *Nkt* is phenotypically similar, however complements *fls* mutants. *Nkt* homozygotes show complete loss of scales, teeth and gill rakers resembling the *fls* phenotype (A–C). Heterozygous *Nkt* zebrafish show an intermediate phenotype of scale loss and patterning defect (arrows) 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.g002

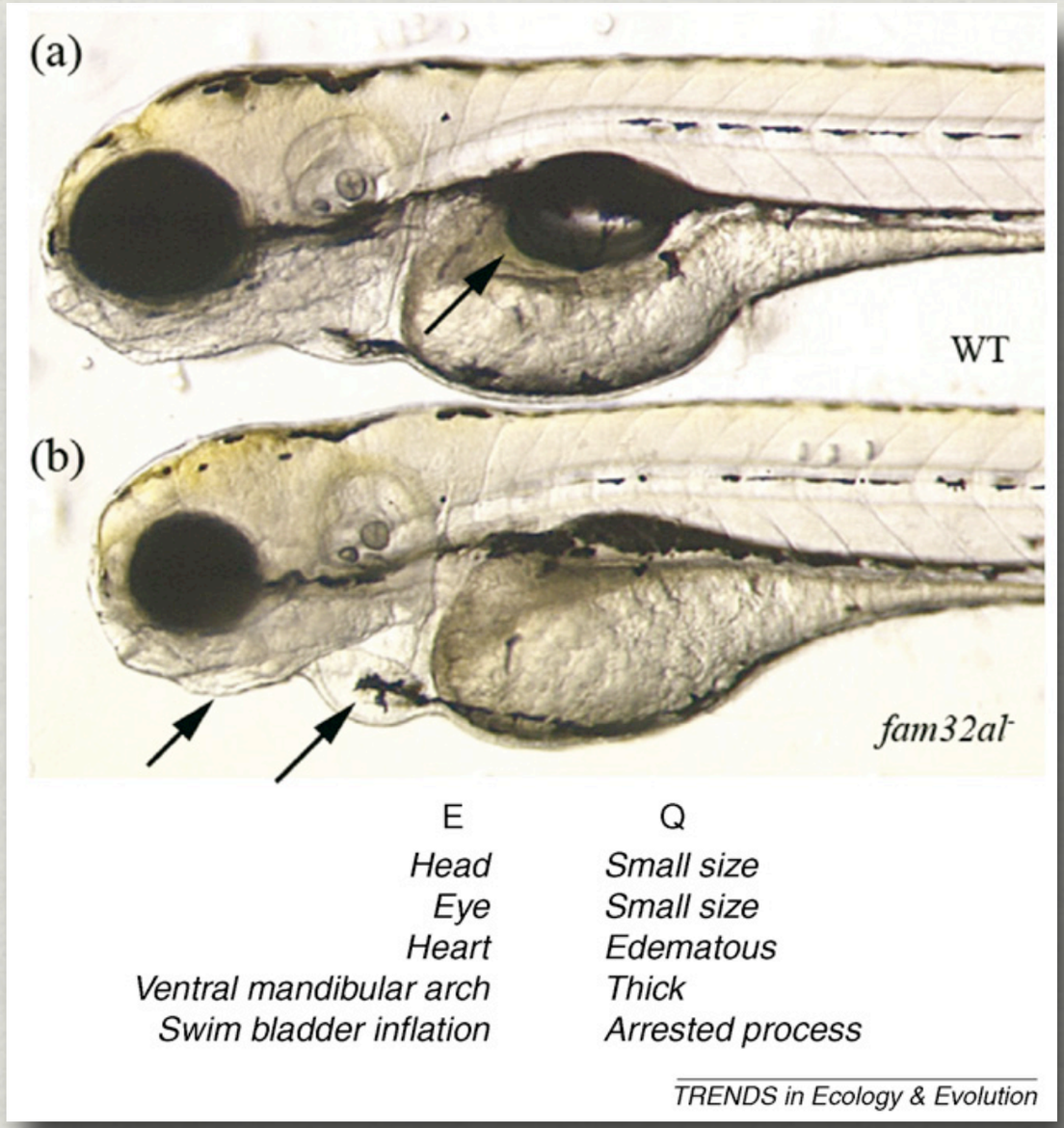
Phenotype annotation

- 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	1 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

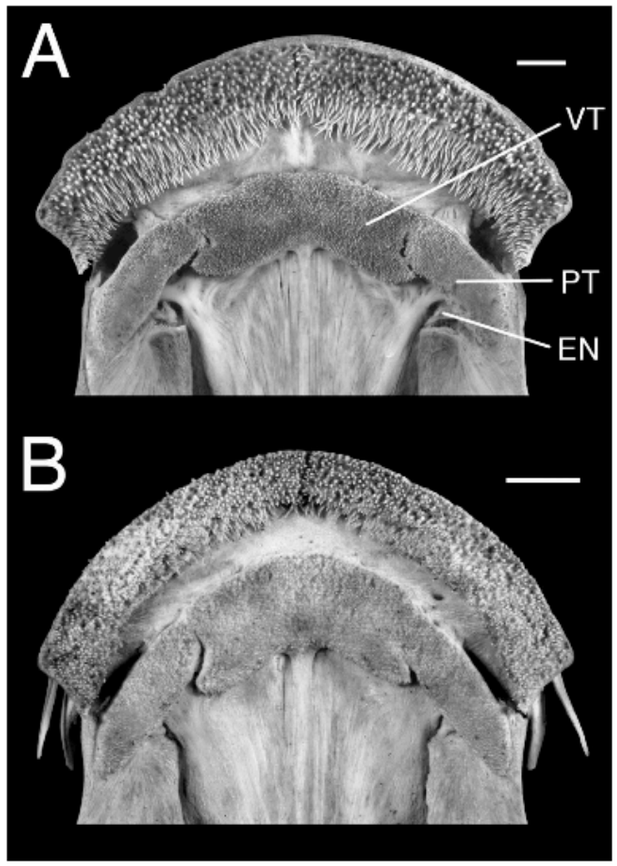
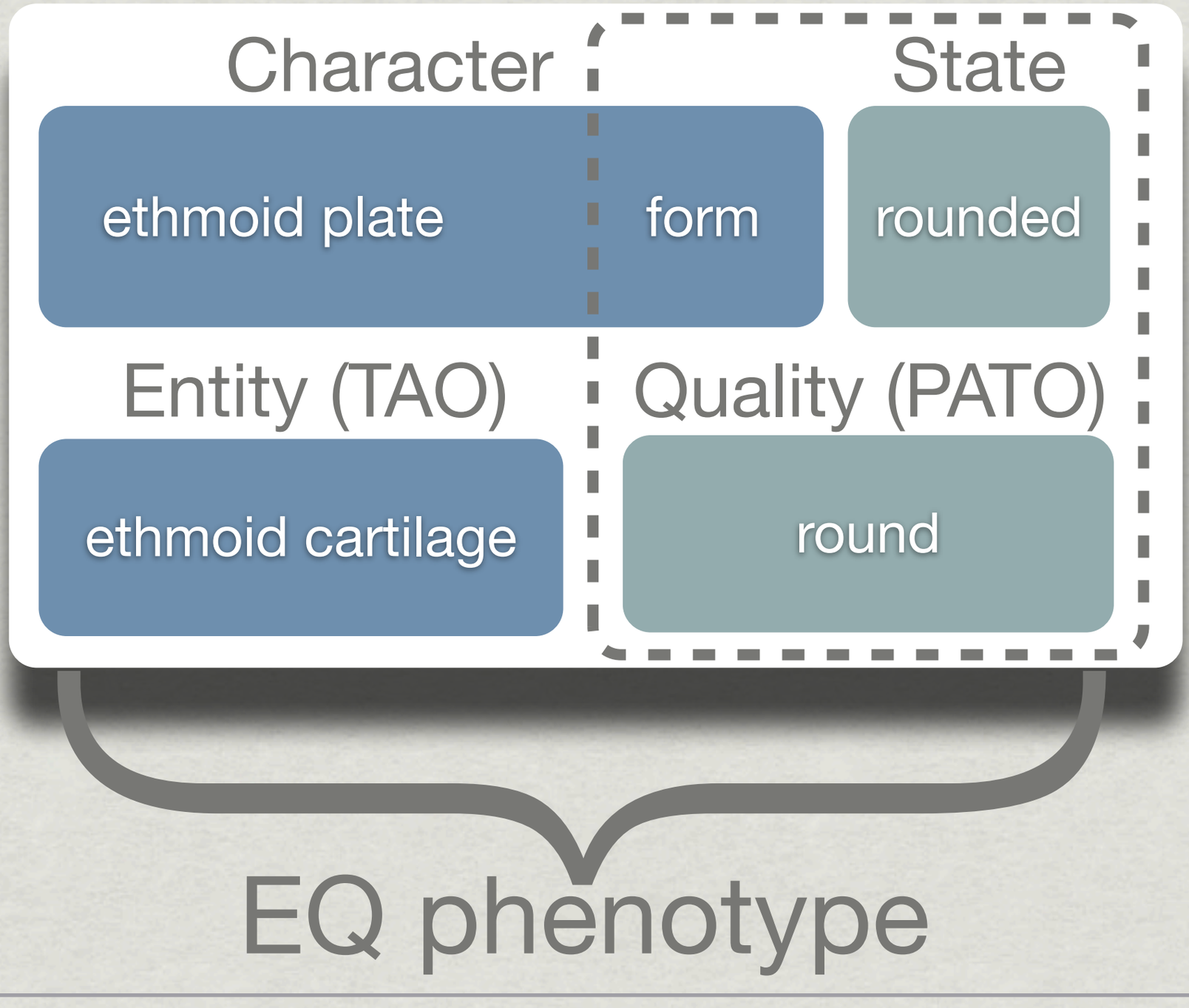


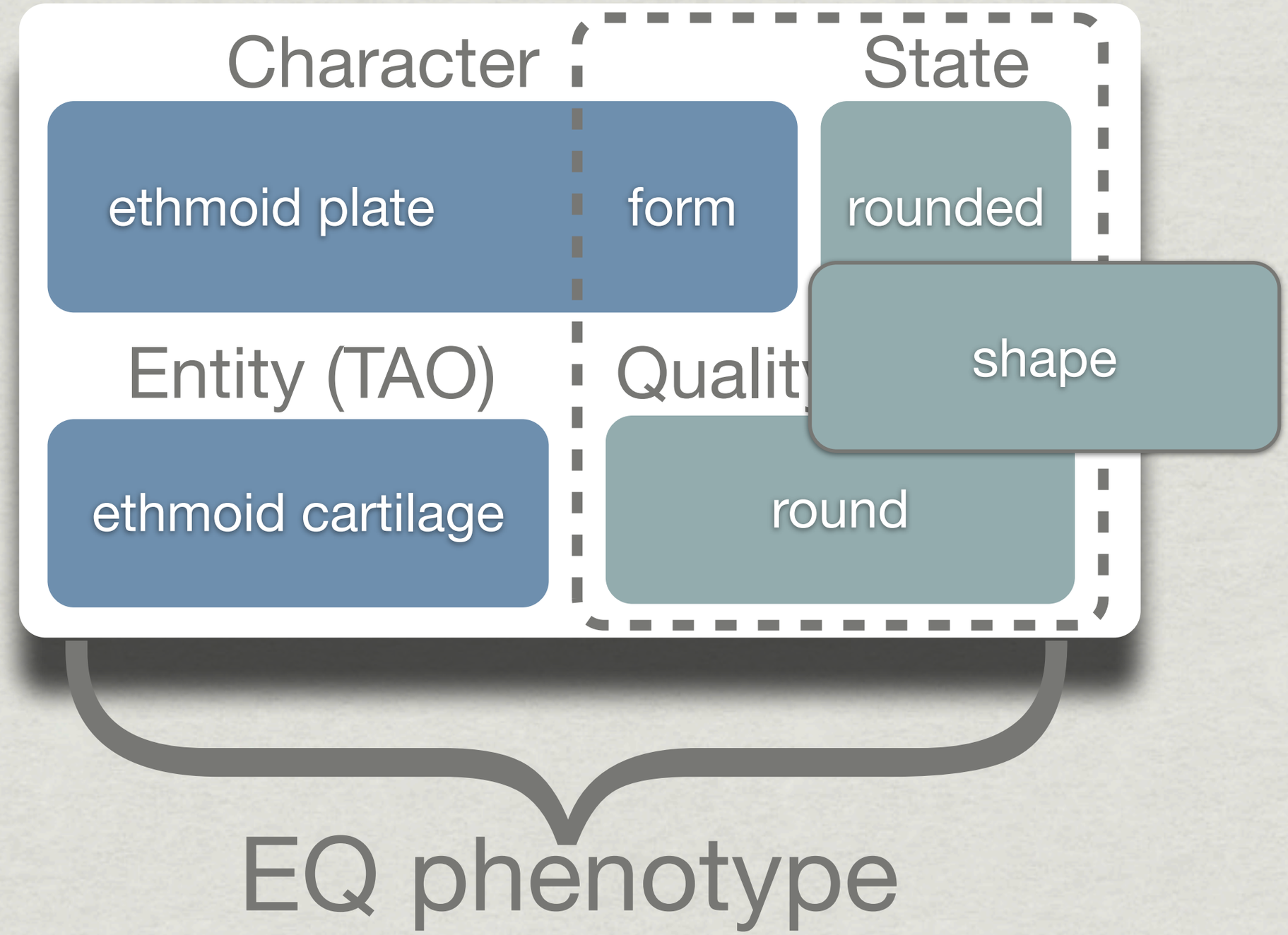
Fig. 7. Dentition of premaxilla and palate in ventral view of (A) *Brachyplatystoma filamentosum*, DU F1052; (B) *Brachyplatystoma capapretum*, MZUSP 53262. Scale bars = 1 cm. PT = pterygoid tooth plate, EN = endopterygoid, VT = vomerine teeth.

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.

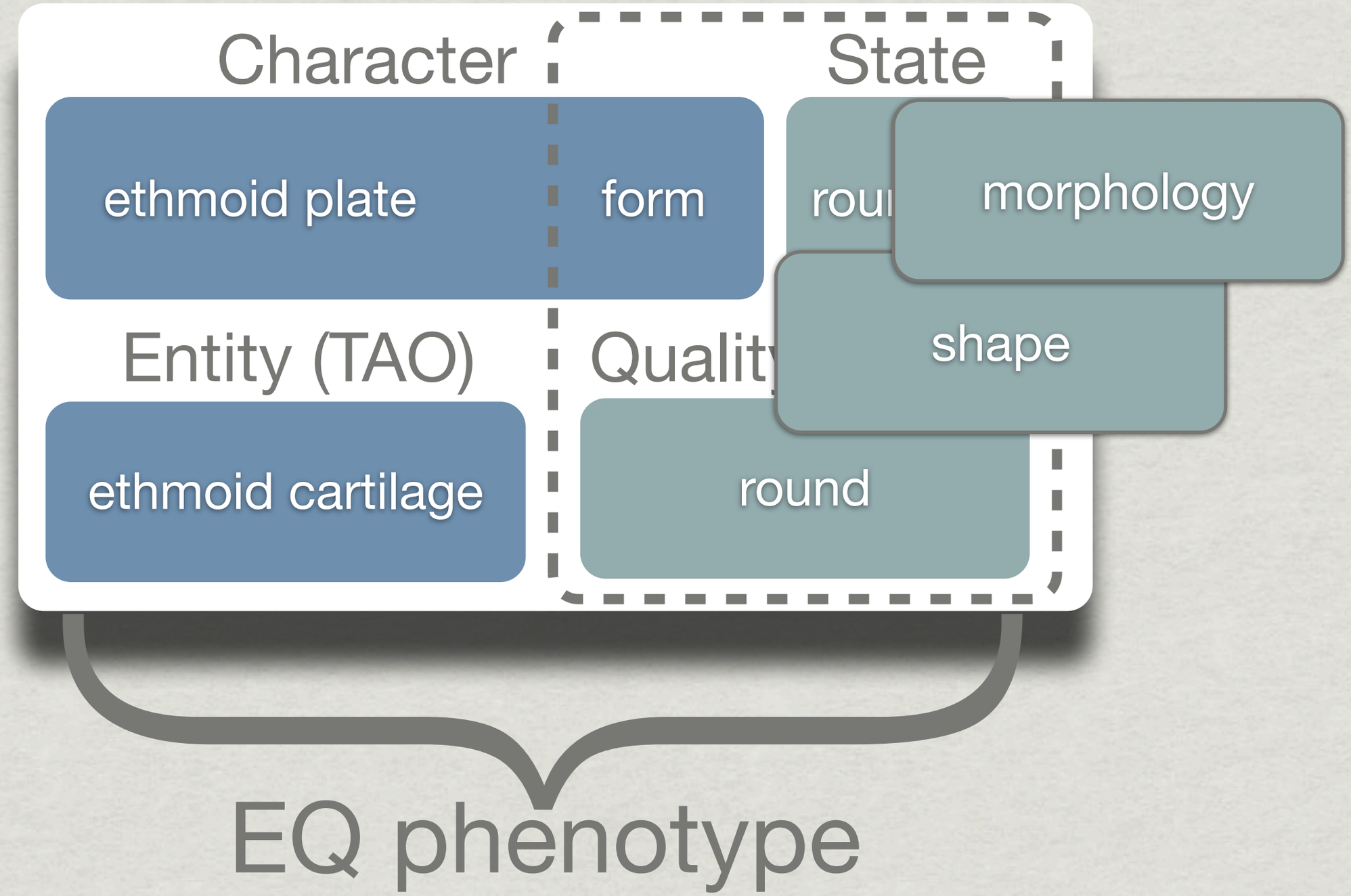
Entity-Quality Model for Evolutionary Phenotypes



Entity-Quality Model for Evolutionary Phenotypes



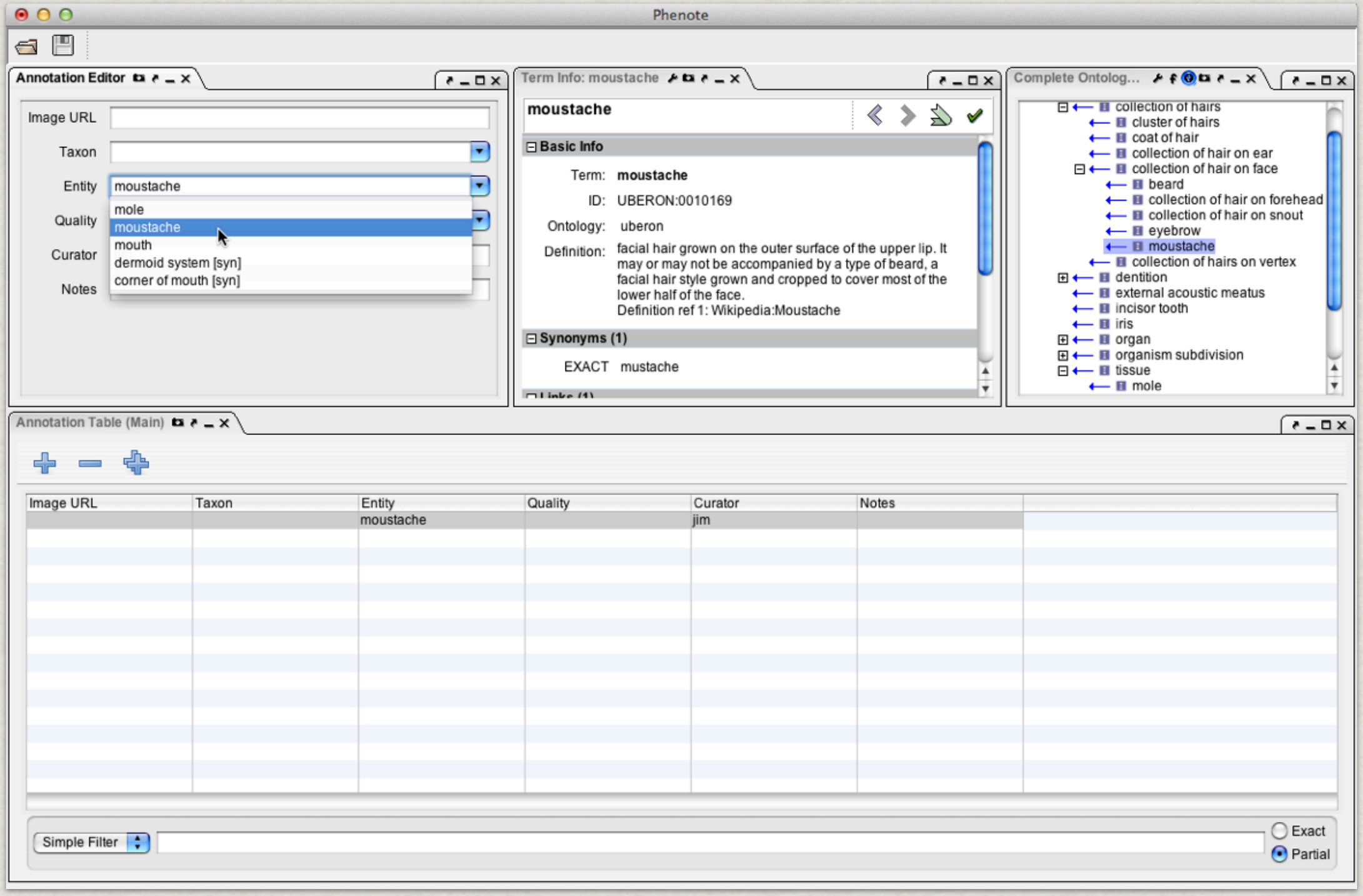
Entity-Quality Model for Evolutionary Phenotypes



Creating annotations

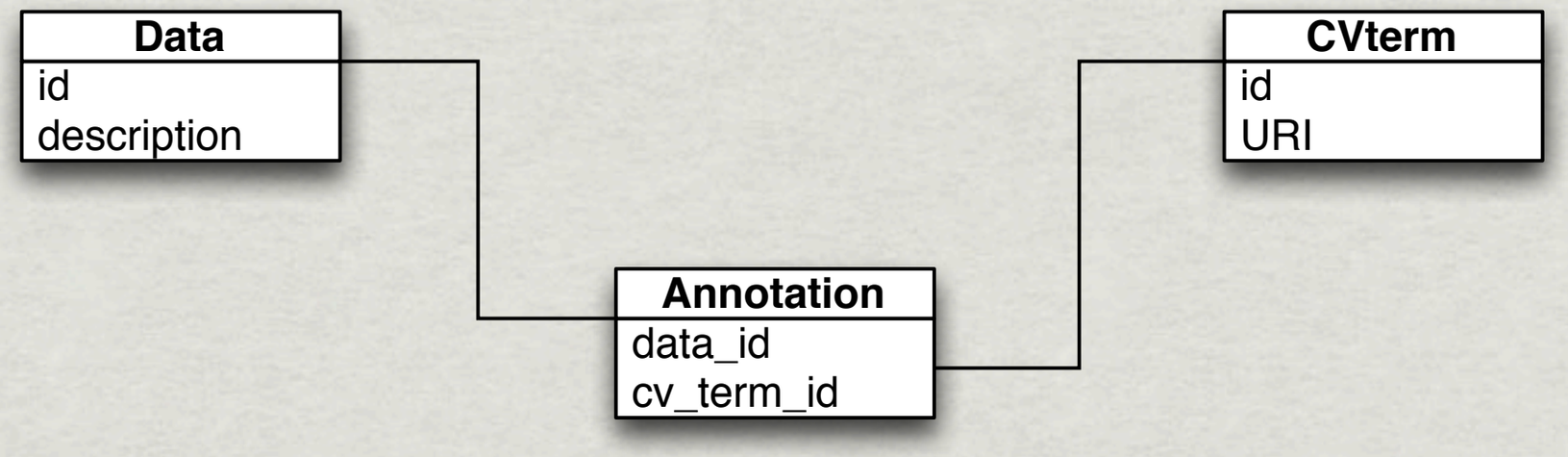
- **Phenote** — generic, configurable table editor
  - loads ontologies, provides term autocomplete
- **Phenex** — specialized for character matrix annotation
- Custom web interfaces
- **Protégé** OWL editor

Phenote


Phenote
Annotation Editor
Image URL
Taxon
Entity
moustache
Quality
Curator
Notes
mole
moustache
mouth
dermoid system [syn]
corner of mouth [syn]
Term Info: moustache
moustache
Basic Info
Term: moustache
ID: UBERON:0010169
Ontology: uberon
Definition: facial hair grown on the outer surface of the upper lip. It may or may not be accompanied by a type of beard, a facial hair style grown and cropped to cover most of the lower half of the face.
Definition ref 1: Wikipedia:Moustache
Synonyms (1)
EXACT mustache
Complete Ontolog...
collection of hairs
cluster of hairs
coat of hair
collection of hair on ear
collection of hair on face
beard
collection of hair on forehead
collection of hair on snout
eyebrow
moustache
collection of hairs on vertex
dentition
external acoustic meatus
incisor tooth
iris
organ
organism subdivision
tissue
mole
Annotation Table (Main)
Image URL
Taxon
Entity
Quality
Curator
Notes
moustache
jim
Simple Filter
Exact
Partial

Using annotations

- Tagging/keyword — provide search for data using ontology hierarchy
  - “Yellow bristles apparent on head”
  - *PATO:yellow, HAO:seta, HAO:head*



- Can expand keyword query using hierarchy-aware middleware

Using annotations

"Yellow bristles apparent on head"

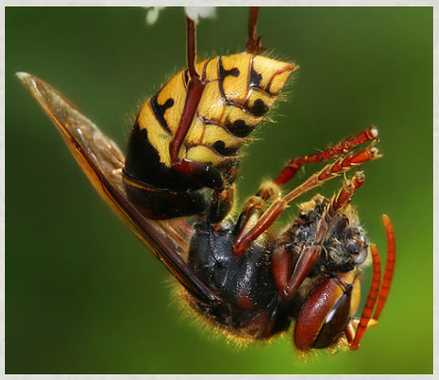
- **E** and **Q** are not always sufficient
  - **Post-composition**: E = "bristle which is part of head"
  - **L**? — locator ("head")
  - **RE**: Related/Dependent Entity — needed for some PATO qualities ("sensitivity toward")

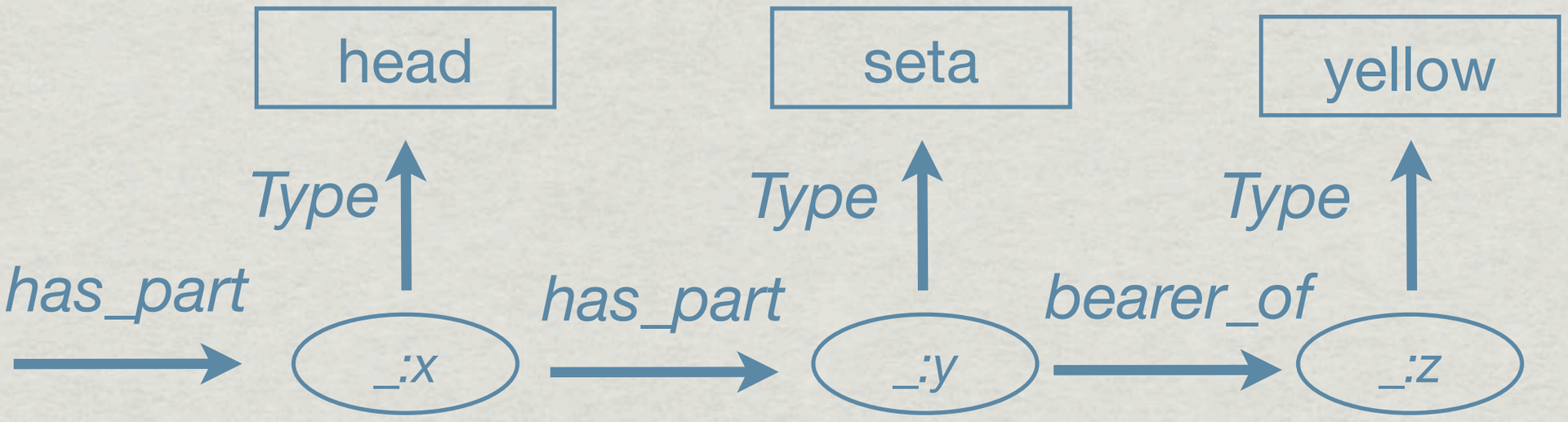
Using annotations

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

Using annotations

- “Yellow bristles apparent on head”
- OWL class expression
  - *has_part* some (**head** and *has_part* some (**seta** and *bearer_of* some **yellow**))

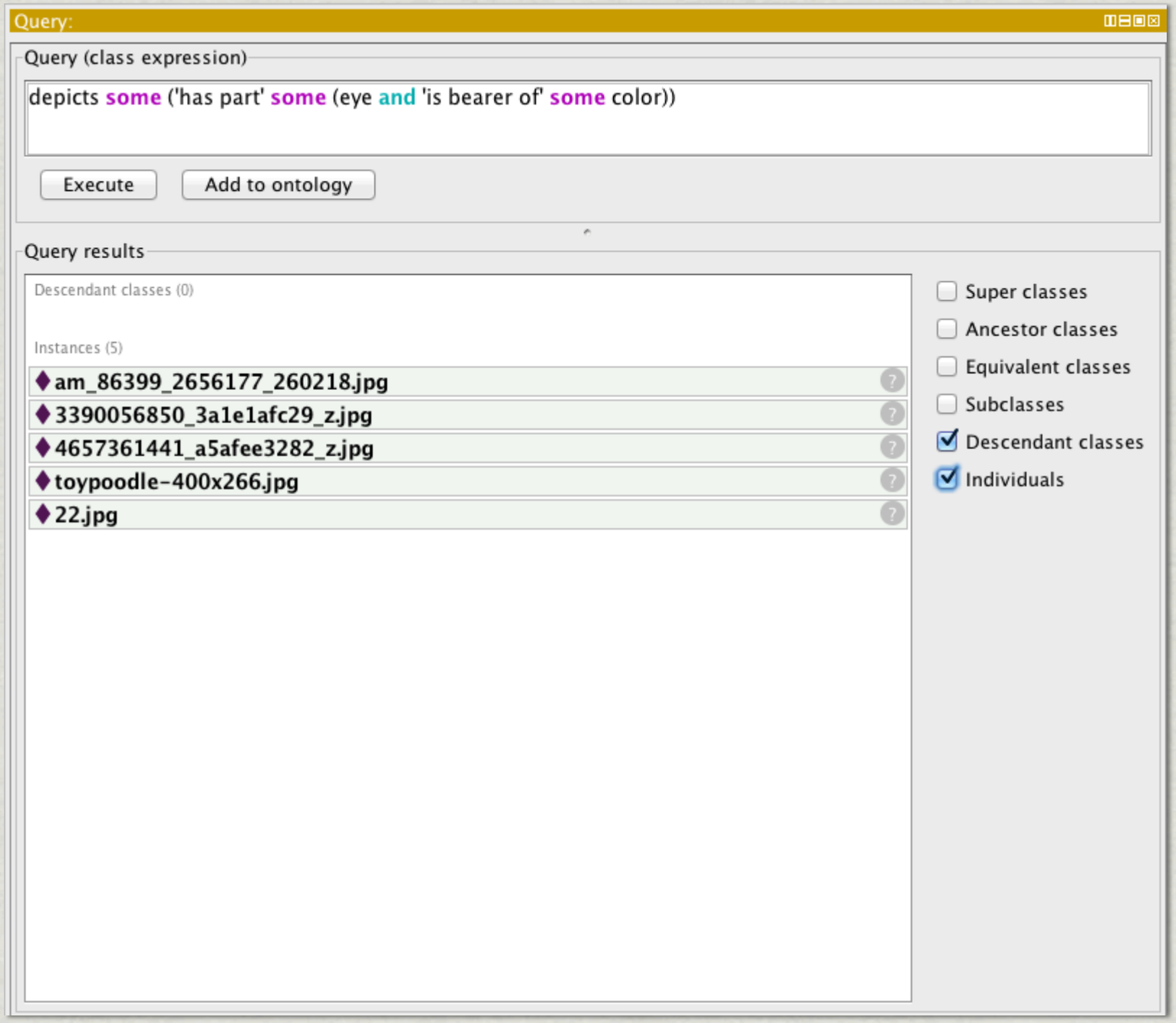




Using annotations

- Can use semantic annotations in reasoner enabled environments
  - OWL-API code, Protégé
    - Reasoners: HermiT, FaCT++, Pellet, ELK
  - Phenoscape KB, Virtual Fly Brain

Protégé


Query:
Query (class expression)
depicts some ('has part' some (eye and 'is bearer of' some color))
Execute
Add to ontology
Query results
Descendant classes (0)
Instances (5)
am_86399_2656177_260218.jpg
3390056850_3a1e1afc29_z.jpg
4657361441_a5afee3282_z.jpg
toypoodle-400x266.jpg
22.jpg
Super classes
Ancestor classes
Equivalent classes
Subclasses
Descendant classes
Individuals

Phenoscape KB

PHENOSCAPE

Site search: Go

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

Feedback

Home | Query taxa | Workspace | About | Acknowledgments

Query for:

Taxa

Taxon is: Add

Any

and

Phenotype is any all : Add

fin • shape [broaden/refine]

including parts

and

head • length [broaden/refine]

including parts

and

Publication is any all : Add

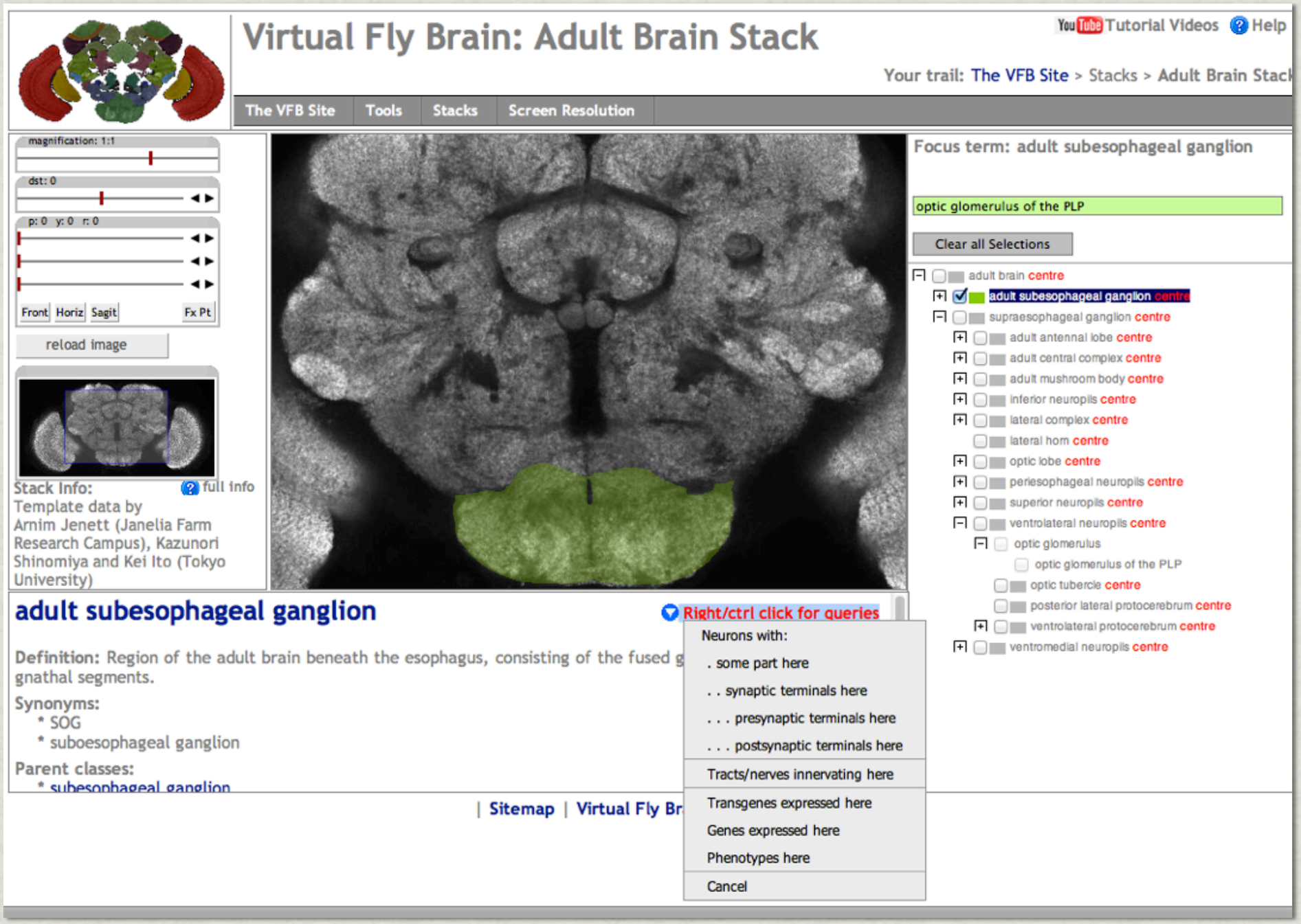
Any

Apply higher taxon annotations to all included species

« Previous Results 1 - 20 of 56 Next »

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)*
Siluriformes	Heptapteridae	*Pimelodella sp. A (Bockmann 1998)*

Virtual Fly Brain



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*