



# *Phenoscape: Using ontologies to link comparative morphology to genes*

**Paula Mabee**

University of South Dakota

Jim Balhoff, Wasila Dahdul, Cartik Kothari, Hilmar Lapp, John Lundberg, Peter Midford, Todd Vision,  
Monte Westerfield



# Phenoscape Knowledgebase (KB)

[kb.phenoscape.org](http://kb.phenoscape.org)

# Comparative morphology



*Cyprinus carpio*



*Pangio anguillaris*



*Nemacheilus fasciatus*



*Catostomus commersoni*



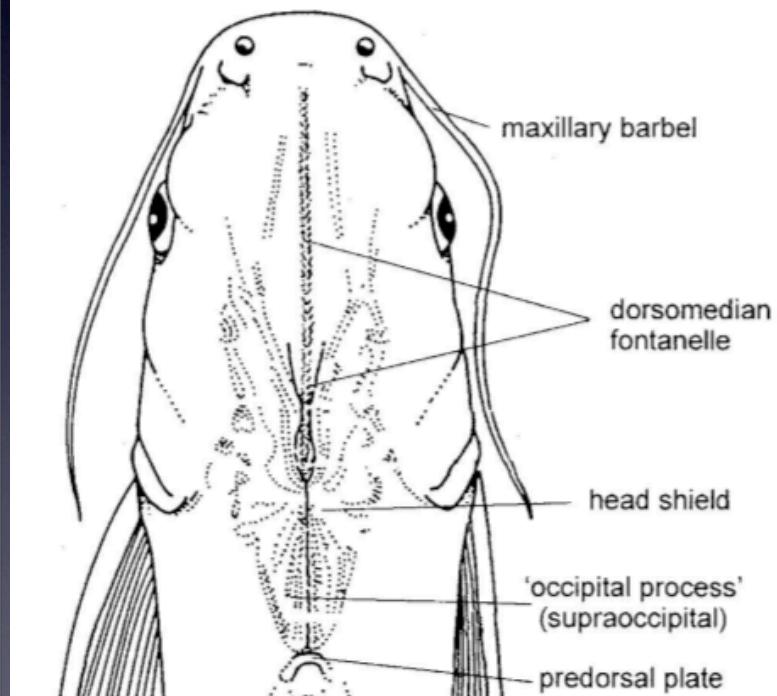
*Gyrinocheilus aymonieri*



*Phenacogrammus interruptus*

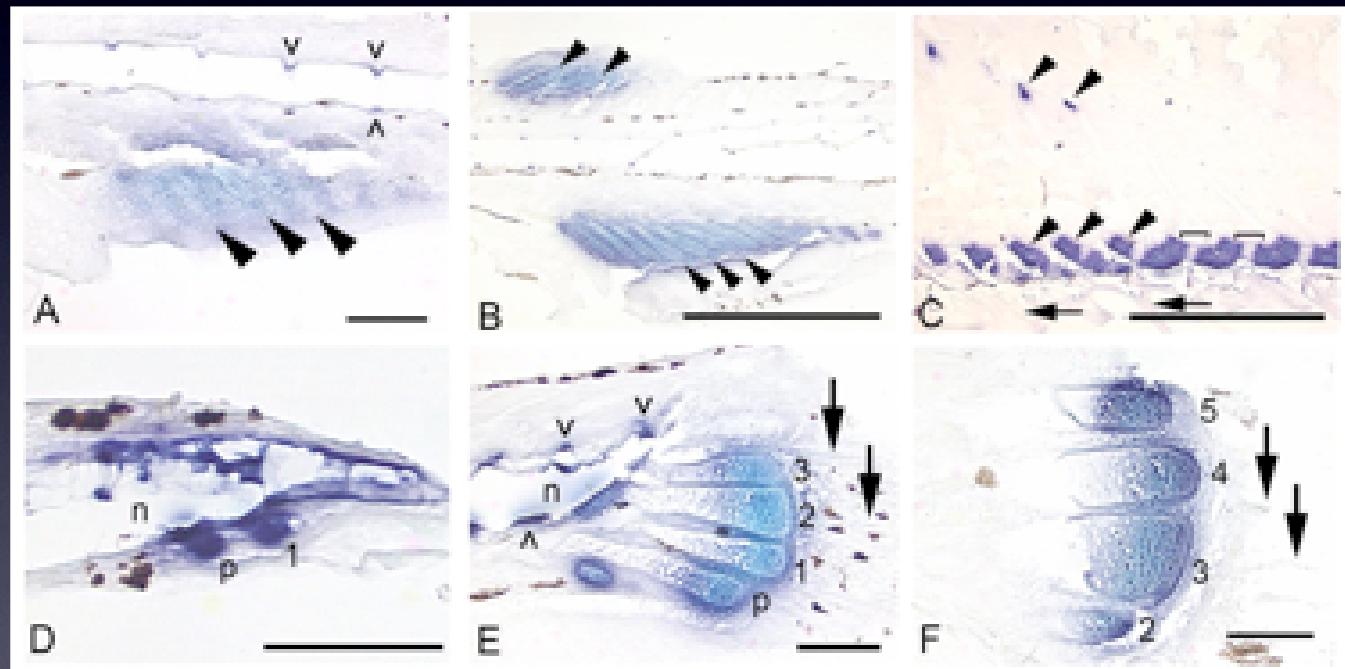
# “Free-text” format not computable across studies

47. *Shape of posterior dorsomedian fontanelle.* An elongate-rectangular posterior fontanelle of most catfishes appears to be the plesiomorphic condition (see also Tilak 1963, 1964, 1965a; Lundberg 1982; Arratia 1987; Grande 1987). An irregularly-shaped or rounded



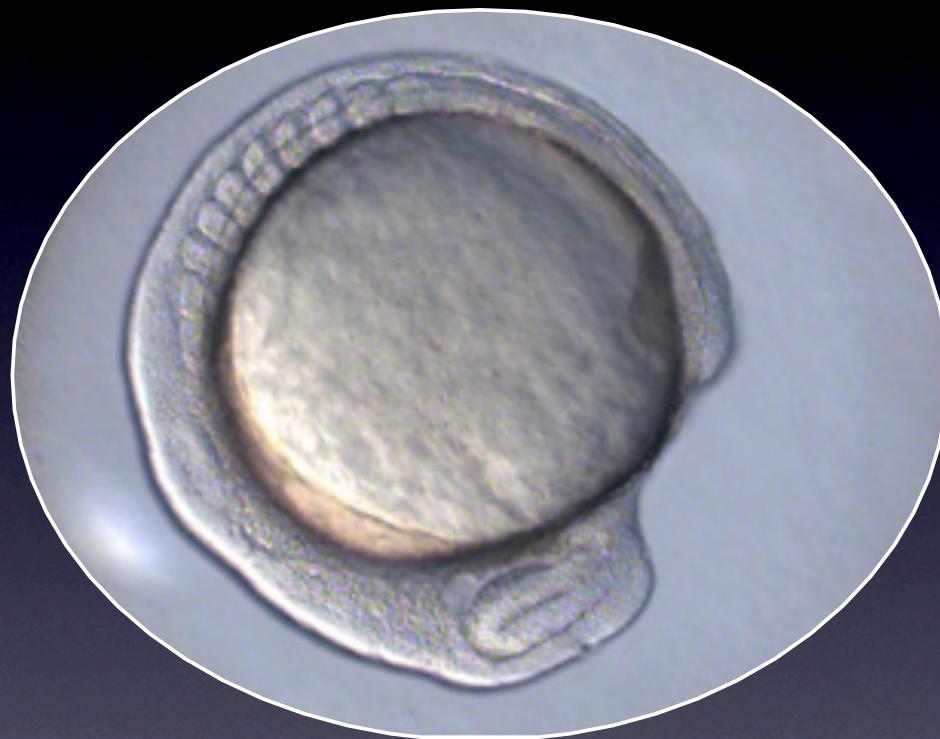
	State 0 = frontal	broad anteriorly and moderately narrow posteriorly	wide posteriorly	deeply excavated	laminar	extensive laminar shelf	evidence of a single median keel	the laminar shelf is very broad	of the transverse processes	I consider this to be plesiomorphic	is plesiomorphic	cover	as exposed transverse process bases and a deep median excavation on the ventral surface	The laminar bone in ariids extends over four to eight vertebra centra	Some ariids possess apomorphic modifications in the laminar shelf, such as depressions (e.g., <i>Guiritinga barbus</i> , <i>Cinetodus froggatti</i> ) or median single keel (e.g., high and acute in <i>Batrachocephalus</i> , <i>Nemapteryx armiger</i> ) or double keel (e.g., <i>Bagre marinus</i> )
1	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
2	?	0	0	?	3	0	?	0	?	0	0	?	1	0	?
3	1	?	?	?	0	?	?	0	1	0	0	2	0	2	0
4	0	?	?	?	0	?	?	0	?	0	0	2	1	2	1
5	0	?	?	?	0	?	?	0	?	0	0	2	1	2	1
6	0	?	?	?	0	?	?	0	1	?	?	2	1	2	0
7	1	?	0	?	2	0	?	0	1	0	?	2	1	2	?
8	1	?	?	?	0	0	?	0	?	0	?	2	1	1	1
9	3	?	?	?	0	?	?	0	1	0	1	2	1	0	2
10	3	?	?	?	0	?	?	0	1	0	?	2	1	0	1
11	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
12	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
13	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
14	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
15	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
16	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
17	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
18	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
19	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
20	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
21	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
22	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
23	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
24	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
25	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
26	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
27	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
28	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
29	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
30	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
31	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
32	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
33	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
34	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
35	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
36	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
37	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
38	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
39	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
40	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
41	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
42	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
43	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
44	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
45	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
46	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
47	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
48	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
49	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
50	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
51	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
52	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
53	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
54	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
55	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
56	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
57	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
58	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
59	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
60	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
61	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
62	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
63	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
64	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
65	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
66	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
67	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
68	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
69	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
70	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
71	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
72	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
73	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
74	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
75	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
76	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
77	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
78	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
79	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
80	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
81	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
82	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
83	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
84	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
85	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
86	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
87	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
88	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
89	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
90	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
91	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
92	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
93	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
94	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
95	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
96	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
97	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
98	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
99	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
100	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
101	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
102	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
103	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
104	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
105	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
106	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
107	2	?	?	?	0	?	?	0	1	0	?	2	1	0	1
108	2	?	?	?	0	?	?	0	1						

# Cannot be synthesized with developmental and genetic data



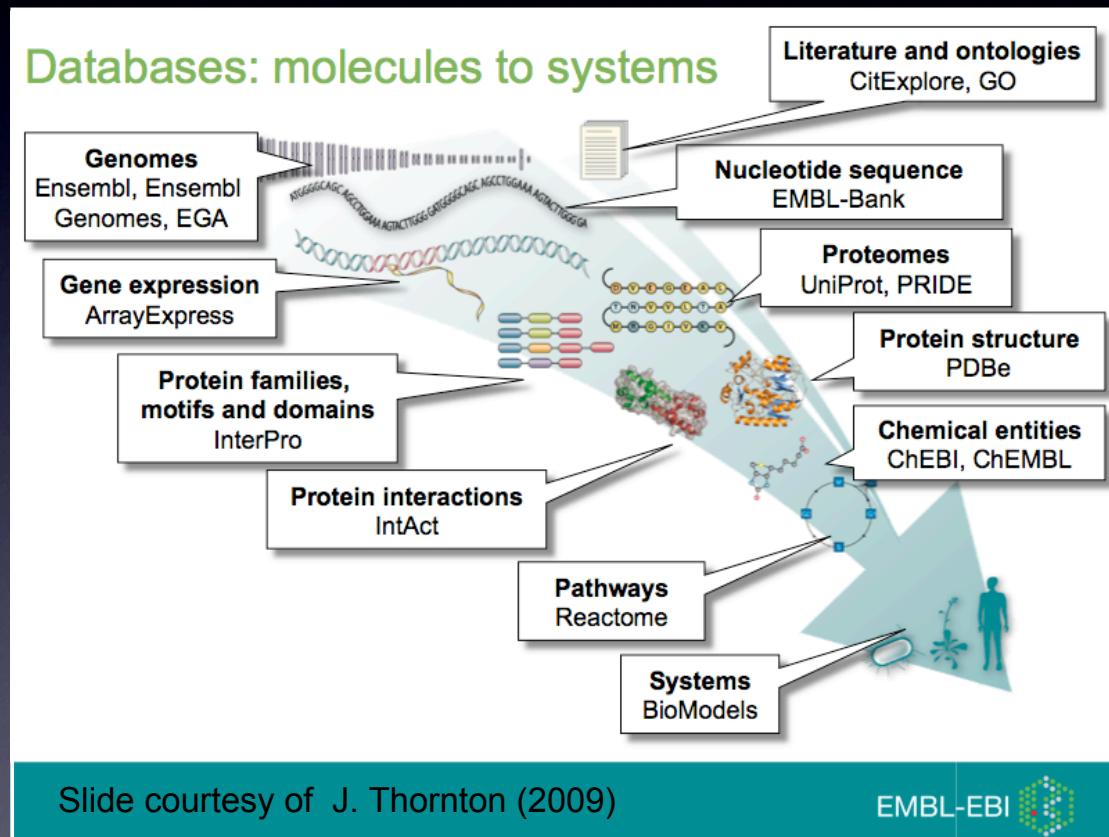
Crotwell & Mabee, 2007

# Morphological differences



have their basis in changes in  
genetic control over development

# >1,170 molecular databases



(Galperin & Cochrane 2008)

# Zebrafish database

- Est. 1992 by Monte Westerfield and Oregon group
- Community resource
- Curate phenotypes of mutants, transgenics, genes, gene expression



The screenshot shows the ZFIN website homepage. The header features the ZFIN logo and navigation links for Research, General Information, and ZIRC. The main content area includes sections for searching genes, mutants, anatomy, publications, people, and laboratories. A sidebar on the right provides links to the Zebrafish International Resource Center, Genomics, Zebrafish Programs, and News. At the bottom right is a small image of a zebrafish embryo.

25 July 2009



← → *ntla*



# Background:

- Mabee (CToL) and Westerfield (ZFIN), 2005-2006, NESCent working group
- Established communication across traditionally separated scientific communities
- Ontology boot camp...
- Interoperability of evolutionary morphology and mutant phenotype data

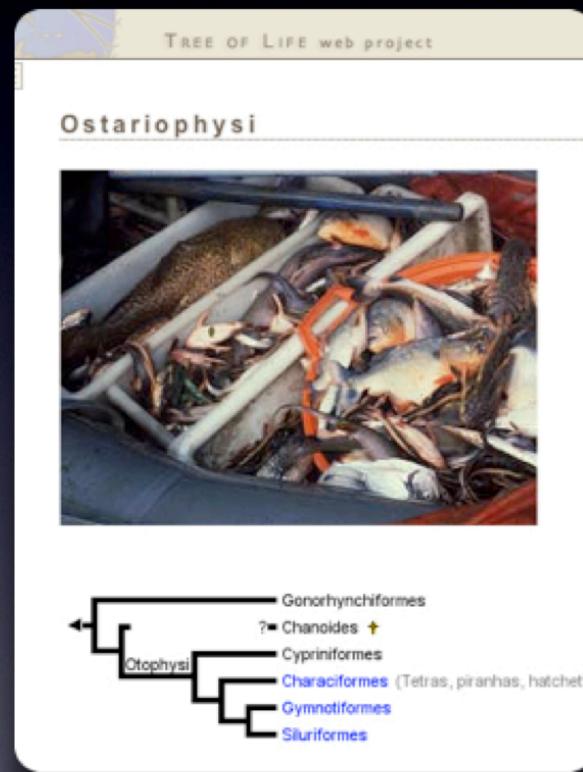


# Phenoscape



2007

- Curated, ontology-based evolutionary morphology database that maps to genetic databases
- Foster devo-evo synthesis
- Data-mining and discovery



25 July 2009; ToL web



# Requirements:

1. Ontologies
2. Curation
3. Database

# I. Ontologies

New:

Zebra  
Onto  
(2196 t

**Teleost Anatomy  
Ontology**  
(2371 terms; 618 skeletal)

**Teleost Taxonomy  
Ontology**  
(36,060 terms;  
38,000 synonyms)

**Taxonomic  
Rank Ontology**  
(8->31 terms)

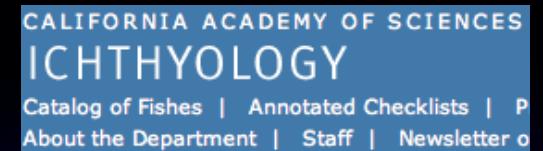
Existing:

**Phenotype and Trait  
Ontology (PATO)**  
(1,075 terms)

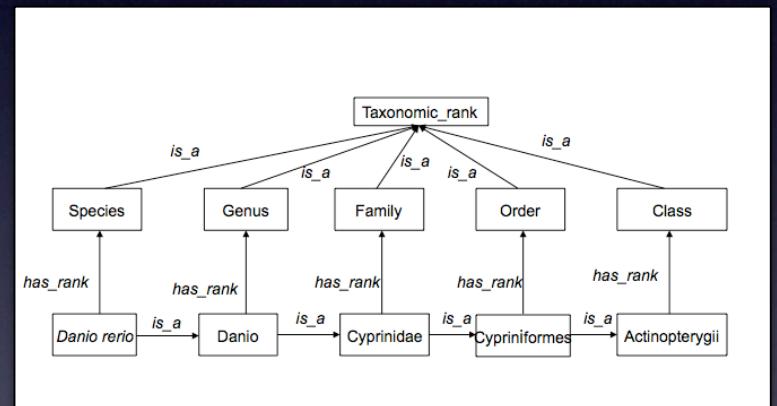
**Spatial Ontology**  
(106 terms)

**Evidence  
Code  
Ontology**

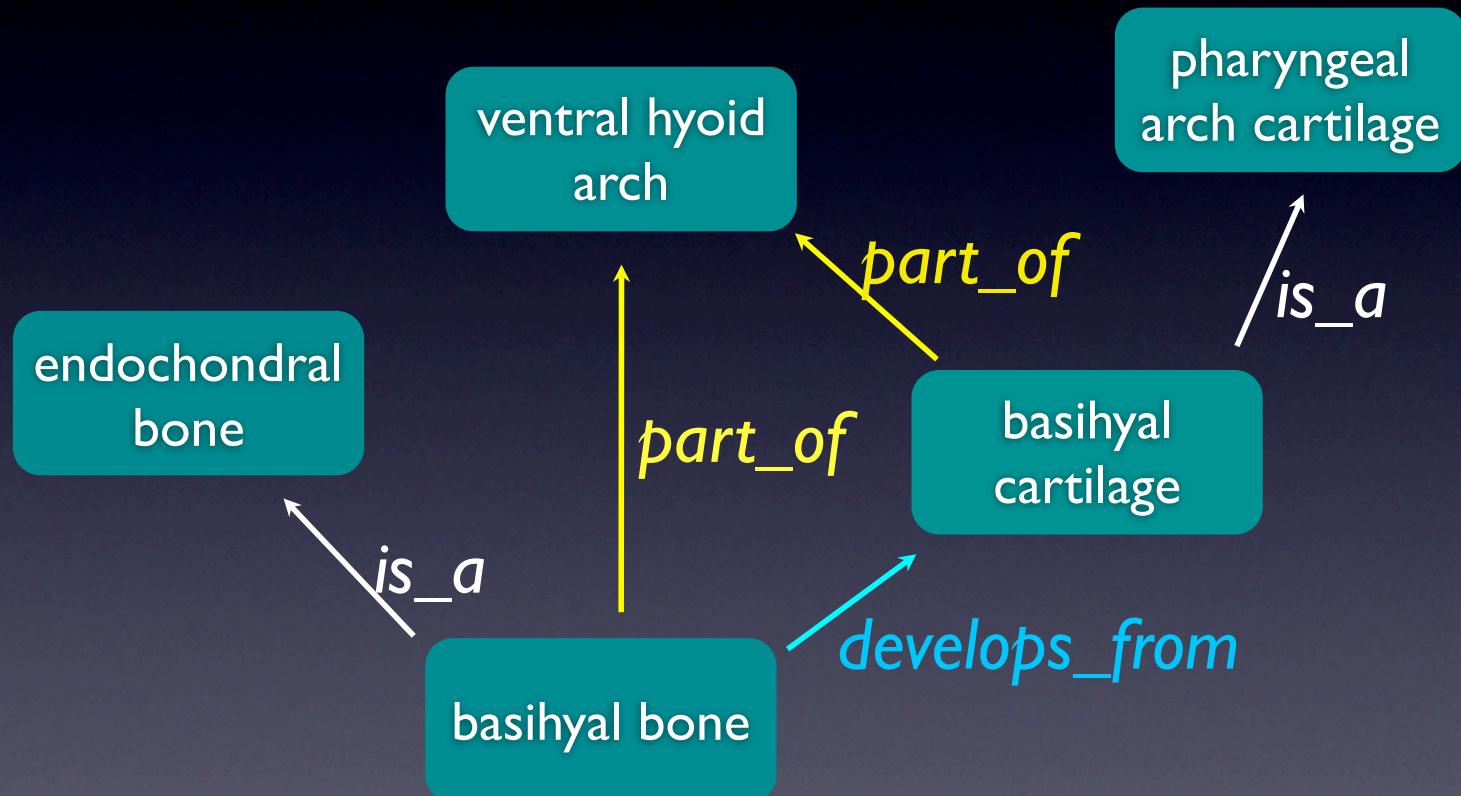
# Teleost Taxonomy Ontology (TTO)



- Based on Eschmeyer's Catalog of Fishes
- Contains 36,060 terms
  - 36,508 are species
  - 5,045 are genera
  - 542 are families
  - 47 are extinct
- Contains 43,122 taxonomic synonyms

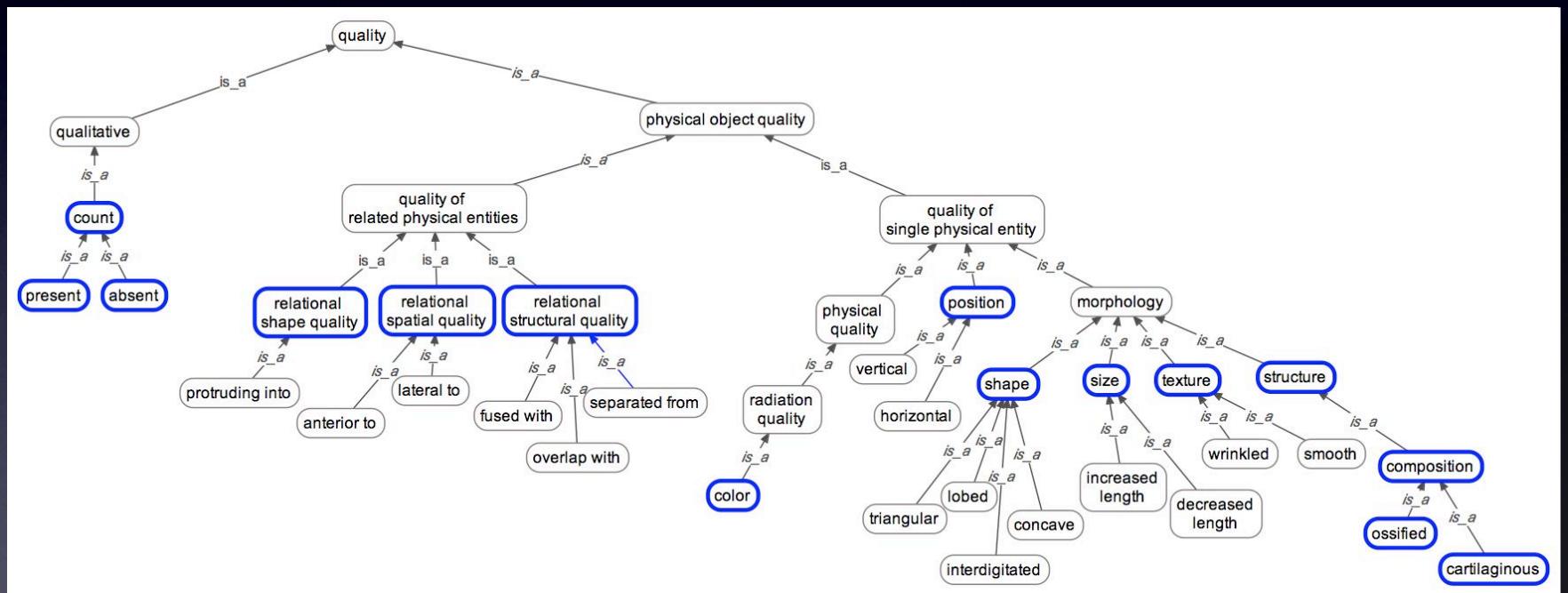


# Teleost Anatomy Ontology (TAO)



Dahdul, W. M., J. G. Lundberg, P. E. Midford, J. P. Balhoff, H. Lapp, T. J. Vision, M. A. Haendel, M. Westerfield, and P. M. Mabee. *in press.* The Teleost Anatomy Ontology: Anatomical representation for the genomics age. *Systematic Biology*.

# Phenotype and Trait Ontology (PATO)

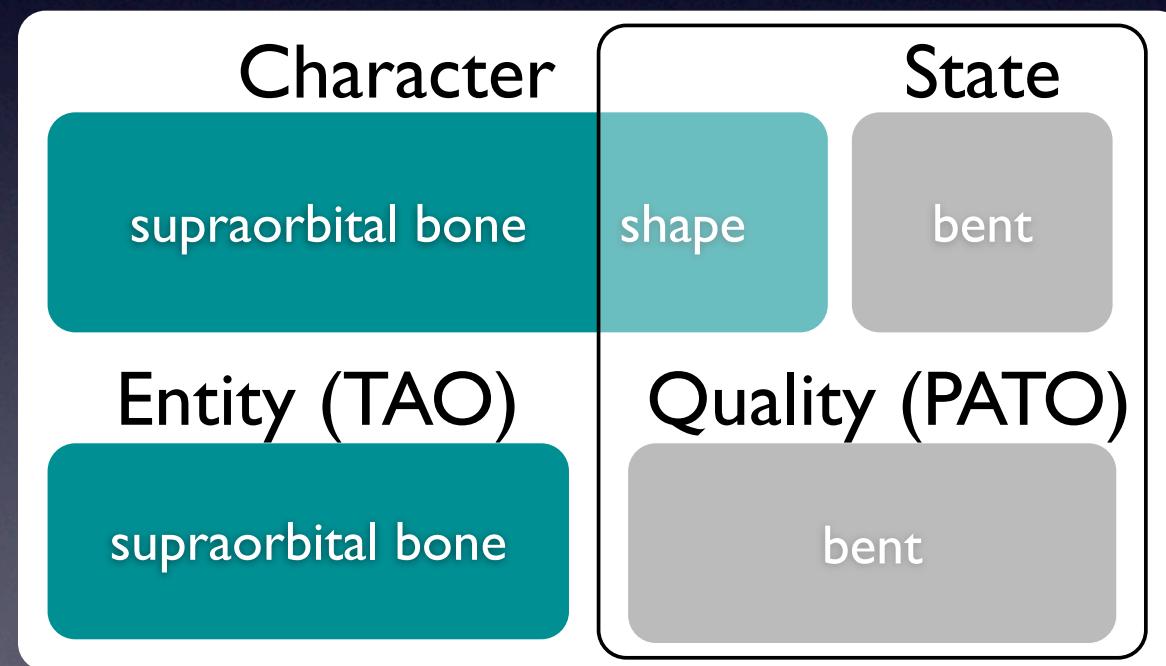


Quality ontology: size, shape, presence-absence, color, etc.

# Phenotype = Entity+Quality

Free text character:

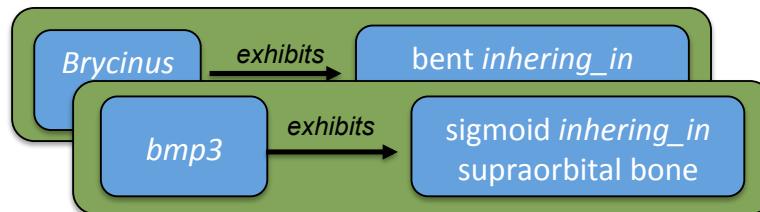
‘Supraorbital bone shape: bent (0) or straight (1)’



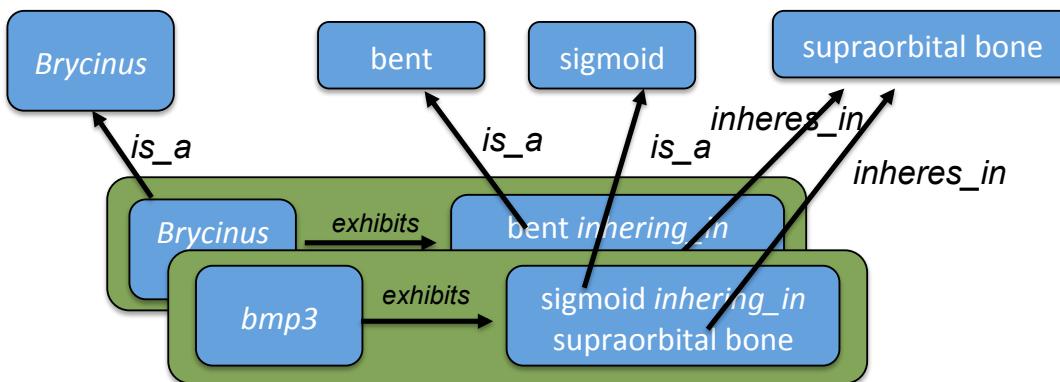
# Taxon + Phenotype = Phenotype assertion



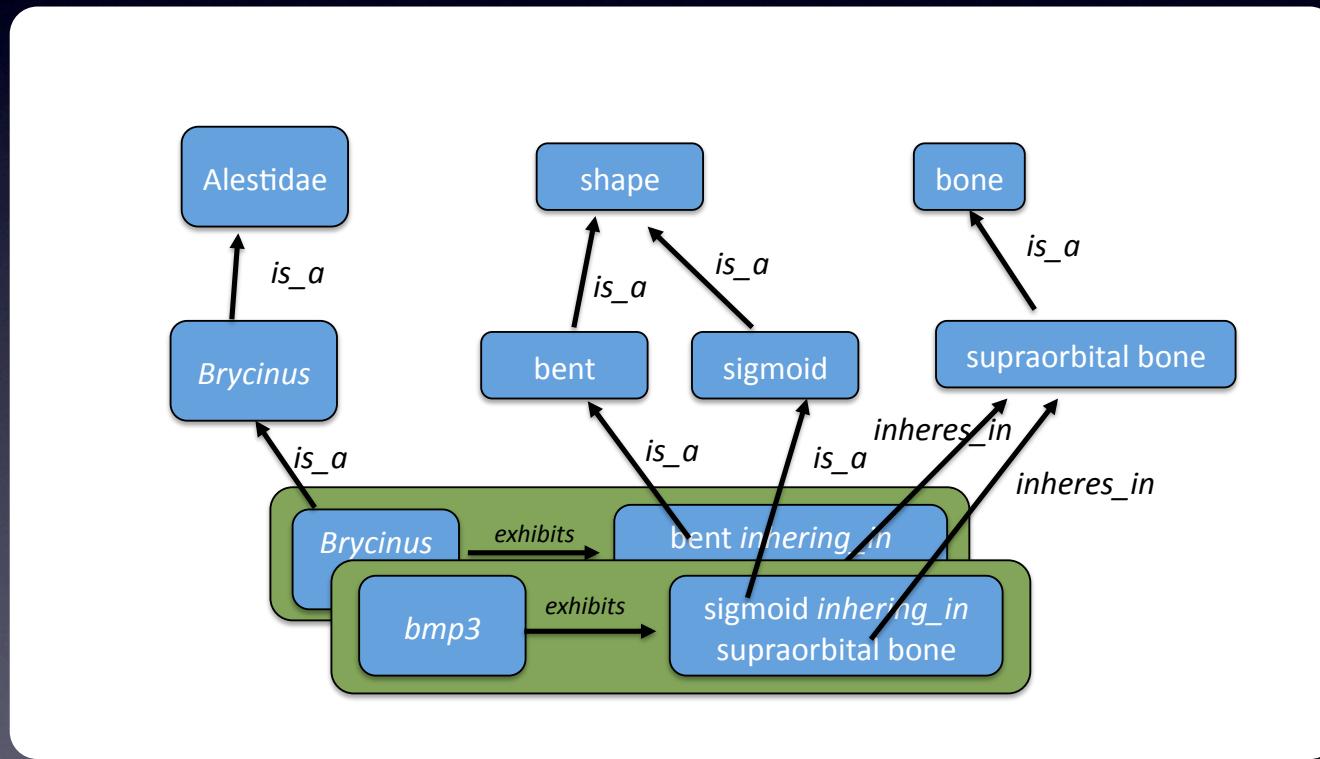
# + Zebrafish phenotype assertions



# Ontologies

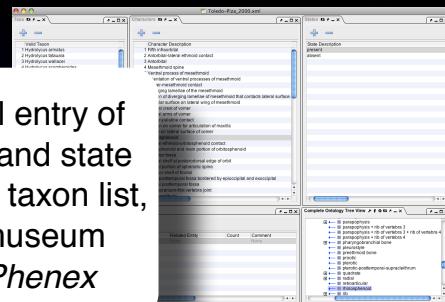


# Data + Ontologies in database

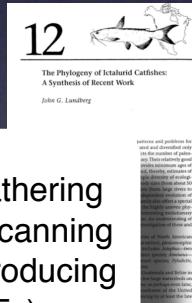


# 2. Curation

2. Student: Manual entry of free text character and state descriptions, matrix, taxon list, specimens and museum numbers using *Phenex*



1. Student: gathering publications (scanning hard copies, producing OCR PDFs)



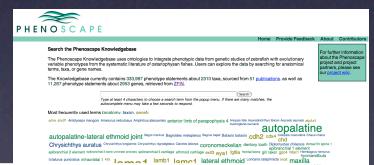
~ 5 person years

**Curators:**  
Wasila Dahdul  
Miles Coburn  
Jeff Engeman  
Terry Grande  
Eric Hilton  
John Lundberg  
Paula Mabee  
Richard Mayden  
Mark Sabaj Pérez

3. Character annotation by experts: Entry of phenotypes and homology assertions using *Phenex*



4. Consistency checks, upload of data to public view of Phenoscape KB



# Free-text characters to phenotypes

86

AMERICAN MUSEUM NOVITATES

NO. 3286

## APPENDIX 1. CHARACTER SUMMARY

1. Fifth infraorbital. 0, well developed, without contact between fourth and sixth infraorbitals; 1, greatly reduced, with posteroventral margin of sixth infraorbital in contact with posterodorsal margin of fourth infraorbital.
2. Antorbital-lateral ethmoid contact. 0, no contact; 1, antorbital contacting ventral wing of lateral ethmoid along its entire lateral edge.
3. Antorbital. 0, flat, platelike, without medial process; 1, with a short medial, vertically aligned process at its posterior edge that extends along posterior surface of ventral wing of lateral ethmoid; 2, with enlarged medial, vertically aligned process at its posterior edge that extends along posterior surface of ventral wing of lateral ethmoid.
4. Mesethmoid spine. 0, conical, or with a dif-
- tilaginous surface at posterior portion of main body of vomer.
14. Portion on vomer for articulation of maxilla. 0, not modified in 1; 1, Presence of a shallow depression on its anterolateral surface where anterior tip of maxilla abuts.
15. Ridge on lateral surface of vomer. 0, absent; 1, present.
16. Rhinosphenoid. 0, present; 1, absent.
17. Lateral ethmoid-orbitosphenoid contact. 0, absent; 1, present.
18. Parasphenoid and main portion of orbitosphenoid. 0, well separated; 1, close to each other.
19. Dilatator fossa. 0, not extending anteriorly on dorsal surface of frontal or if so, only to dorsoposterior edge of orbit; 1, highly developed, extending anteriorly on dorsal surface of frontal beyond dorsoposterior edge of orbit.

(Toledo-Piza 2000)

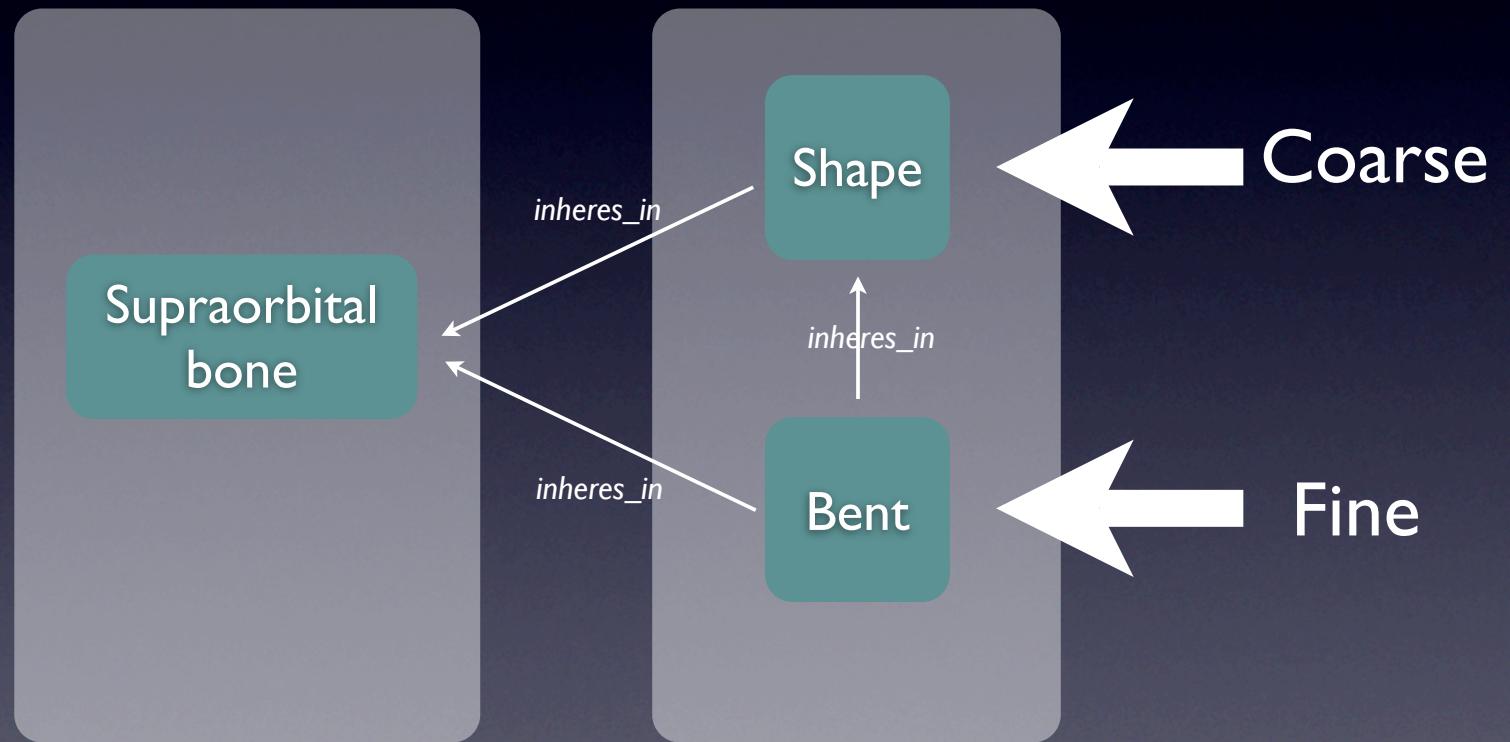
# Phenex software for curation

The screenshot displays the Phenex software interface with five main windows:

- Taxa**: A list of valid taxa, including species from Hydrolycus, Cynodon, Roestes, Gilbertolus, Acestorhynchus, and others, along with two entries for Toledo-Piza 2000.
- Characters**: A list of character descriptions numbered 1 to 26. Character 16, "Rhinophenoid", is highlighted with a yellow background.
- States**: A list of state descriptions: present and absent.
- Phenotypes**: A table showing phenotypic data for the rhinophenoid entity. The entity is listed as "rhinophenoid" with quality "present".
- Complete Ontology Tree View**: A tree view of biological structures, starting with "paraphysis" and branching into various skeletal elements like pharyngobranchial bone, pleurostyle, preethmoid bone, prootic, pterotic, pterotic-posttemporal-supracleithrum, quadrate, radial, retroarticular, rhinosphen, and rib.

In the bottom right corner, there is a logo for "nexml" with the tagline "phylogenetic data in xml".

# Curated to a coarse level



Entity from TAO

Quality from PATO

# Curate homology assertions

- Type of annotation supported by evidence (e.g., topological, morphological, developmental similarity)
- Enables user to view homology assertions, examine evidence

# Example: Weberian apparatus

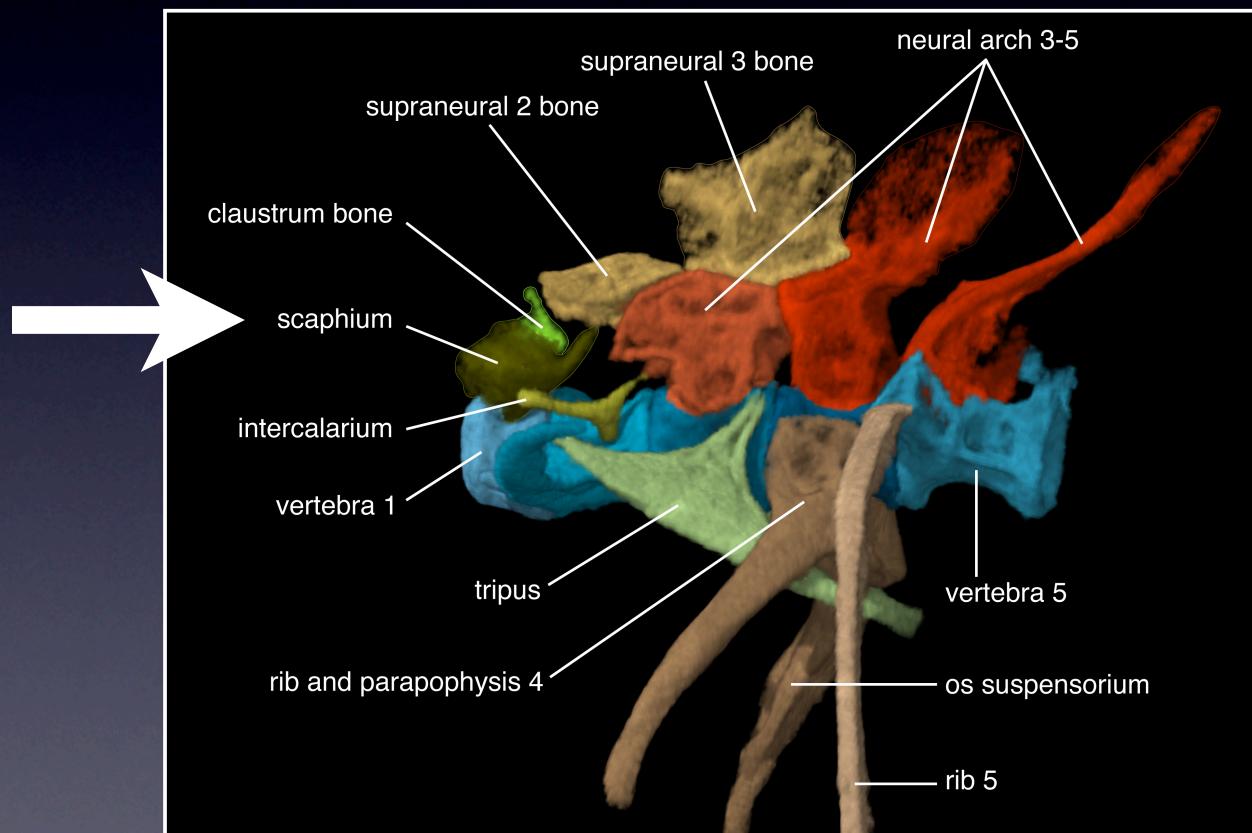


image by Kyle Luckenbill, ANSP

# Curate homology assertions and evidence



Entity 1	Taxon 1	Relationship	Entity 2	Taxon 2	Evidence	Reference(s)
scaphium	Otophysi	homologous_to	neural arch 1	Teleostei	IDS, IMS, IPS	(Fink and Fink, 1981; Rosen and Greenwood, 1970)
intercalarium	Otophysi	homologous_to	neural arch 2 (ventral portion)	Teleostei	IDS, IMS, IPS	(Rosen and Greenwood, 1970)
intercalarium	Otophysi	homologous_to	neural arch 2	Teleostei	NAS	(Fink and Fink, 1981)
intercalarium	Otophysi	homologous_to	neural arch 2	Teleostei	IMS	(Hora, 1922)
intercalarium	Otophysi	homologous_to	rib of vertebra 2	Teleostei	TAS	(Hora 1922)
tripus	Otophysi	homologous_to	parapophysis + rib of vertebra 3	Teleostei	IDS, IMS, IPS	(Fink and Fink, 1981; Rosen and Greenwood, 1970)

Dahdul et al., in press. ‘The Teleost Anatomy Ontology: Anatomical representation for the genomics age.’ *Systematic Biology*

# Homology online...

The screenshot displays the Phenoscape website interface. At the top, there is a logo with three wavy lines and the text "PHENOSCAPE". A search bar at the top right contains the placeholder "New search (anatomy, taxon, gene):" with a "Go" button. Below the header, a navigation menu includes "Home", "Provide Feedback", "About", and "Contributors".

The main content area features a title "anatomical term: scaphium". To the left, a sidebar contains a "Properties" section with ID TAO:0000429, Ontology Teleost Anatomy, and Synonyms Weberian ossicle 2, second Weberian ossicle. It also lists relationships: is a type of Weberian ossicle, neural arch; is part of vertebra 1; and possible homologs neural arch 1 in Teleostei. A "More information" section links to NCBO Bioportal.

The central part of the page shows "Phenotypes" for the term scaphium. It lists five entries:

- Anatomy: scaphium ([relational structure](#), neurocranium)
- scaphium ([relational spatial](#), relational spatial)
- scaphium ([shape](#))
- scaphium ([size](#))
- scaphium ([count](#): absent)

A modal window titled "Source Data" provides detailed information for each entry. For the first entry, it states: "scaphium in Otophysi *homologous to* neural arch 1 in Teleostei". Evidence is inferred from developmental similarity, and citations include Rosen and Greenwood 1970; Fink and Fink 1981. Similar entries are shown for the other four phenotypes.

To the right, a table summarizes Zebrafish Genes and Taxa:

Zebrafish Genes	Taxa
0	<a href="#">2</a>
0	<a href="#">281</a>
0	<a href="#">273</a>
0	<a href="#">54</a>
0	<a href="#">23</a>

# Curated 4,208 characters in 2,310 species from 51 papers

TAXON	# PAPERS	# SPECIES	# CHARACTERS
Cypriniforms	8	676	794
Siluriforms	20	1724	2,110
Characiforms	10	754	1,156
Gymnotiforms	1	116	231
Gonorynchiforms	3	41	467
Clupeiforms	5	200	439
Euteleosts	3	145	582
Total	51	3,656	5,779
<b>TOTAL ANNOTATED</b>	<b>51</b>	<b>2310</b>	<b>4208</b>

# Phenoscape KB

333,987 evolutionary  
phenotype assertions

11,267 phenotype  
statements about  
2,953 genes

The screenshot shows the Phenoscape Knowledgebase homepage. At the top is the Phenoscape logo with three green wavy lines above the word "PHENOSCAPE". Below the logo is a navigation bar with links for "Search", "Provide Feedback", and "About". A welcome message says "Welcome to the Phenoscape phenotype database for ostariophyan fishes." Below this is a search bar with the placeholder "Begin typing to choose a search term from the popup menu." A "Search" button is next to it. Below the search bar is a list of commonly used terms starting with "ache ahct1 Amblyceps mangensis Ameliurus nebulosus Amphilius atesuensis anterior limb of parapophysis 4". The word "process" is prominently displayed in large green letters across the bottom of the page.

Welcome to the Phenoscape phenotype database for ostariophyan fishes.

Search the Phenoscape Knowledgebase

Begin typing to choose a search term from the popup menu.

Search

Or, try one of the most commonly used terms:

ache ahct1 Amblyceps mangensis Ameliurus nebulosus Amphilius atesuensis anterior limb of parapophysis 4 Ariopsis felis Aspredinichthys tibicen Aspredo aspredo atp1a1 Austroglanis barnardi **autopalatine** cdh2 cdx4 Cetopsis coeculenta Chaca chaca chd Chrysichthys auratus Chrysichthys longipinnis Chrysichthys nigrodigitatus Clarotes laticeps coronomeckelian dentary tooth Diplomyxtes chilensis dorsal fin spine 1 epibranchial 1 element epibranchial 2 element epibranchial 3 bone uncinate process epibranchial 5 cartilage ethmoid cartilage extl3 eya1 fgf8a gill bone gill raker gpc4 hdacl Hemibagrus nemurus hommandibula Ictalurus punctatus infraorbital 1 kita **lama1** lamb1 **lamic1** lateral ethmoid Loricaria cataphracta lrc6 maxilla mesethmoid bone metapterygoid mib mycbp2 Mystus nigriceps nld2 Nematogenys inermis opercle os suspensorium Pangasianodon hypophthalmus Parakysis verrucosus **parapophysis** pax2a posterior process of basipterygium pou5f1 premaxilla premaxillary tooth Pseudolaguvia shawi Rhabdalestes septentrionalis Rhamdia laticauda Rita rita scale Schilbe mystus sensory canal shha smo tdo tp53 transcapular ligament trpm7 vangl2 vomer wnt11 wnt5b

**process**

25 July 2008

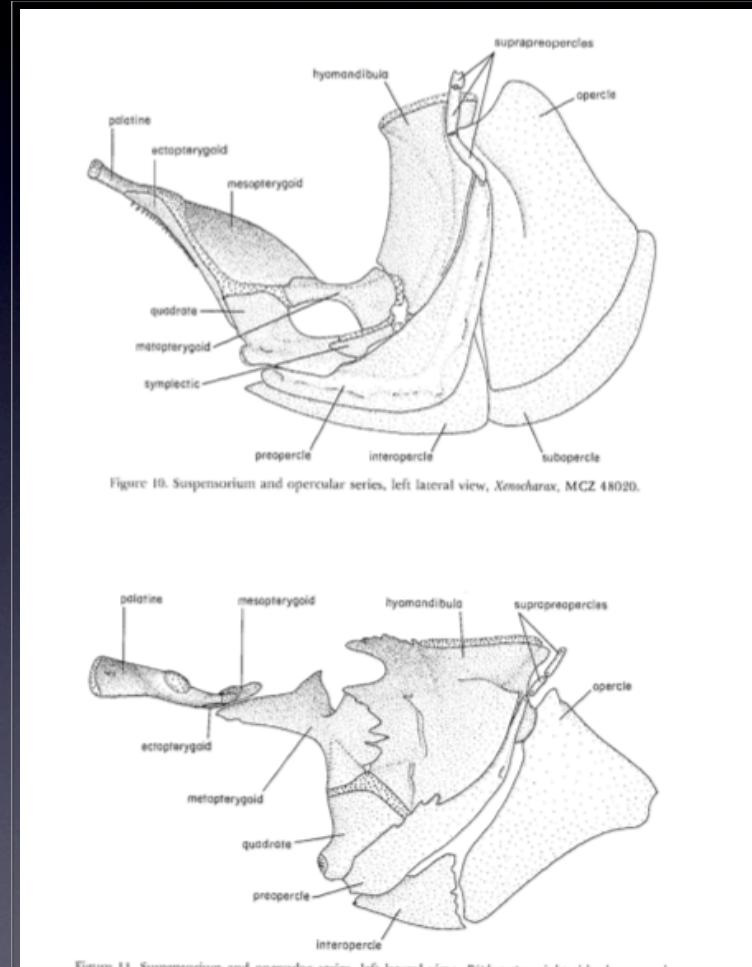


# Basic overview of data

# 1 Character state: 1 Phenotype

Character 36 (Fink & Fink, 1981). “In siluriforms the **opercle** is approximately **triangular** in shape (0) rather than approximately rectangular (1) as in other ostariophysans and primitive teleosts.”

**Phenotype:**  
E: **opercle**; Q: **triangular**



(Fink and Fink 1981)

# 1 Character state: 2 Phenotypes

Character 80 (Zanata & Vari, 2005). “Form and area of attachment of **primordial ligament**: (0) ligament relatively **narrow** and **attaching to** posteromedial portion of **ascending process of maxilla**.....

## 2 Phenotypes:

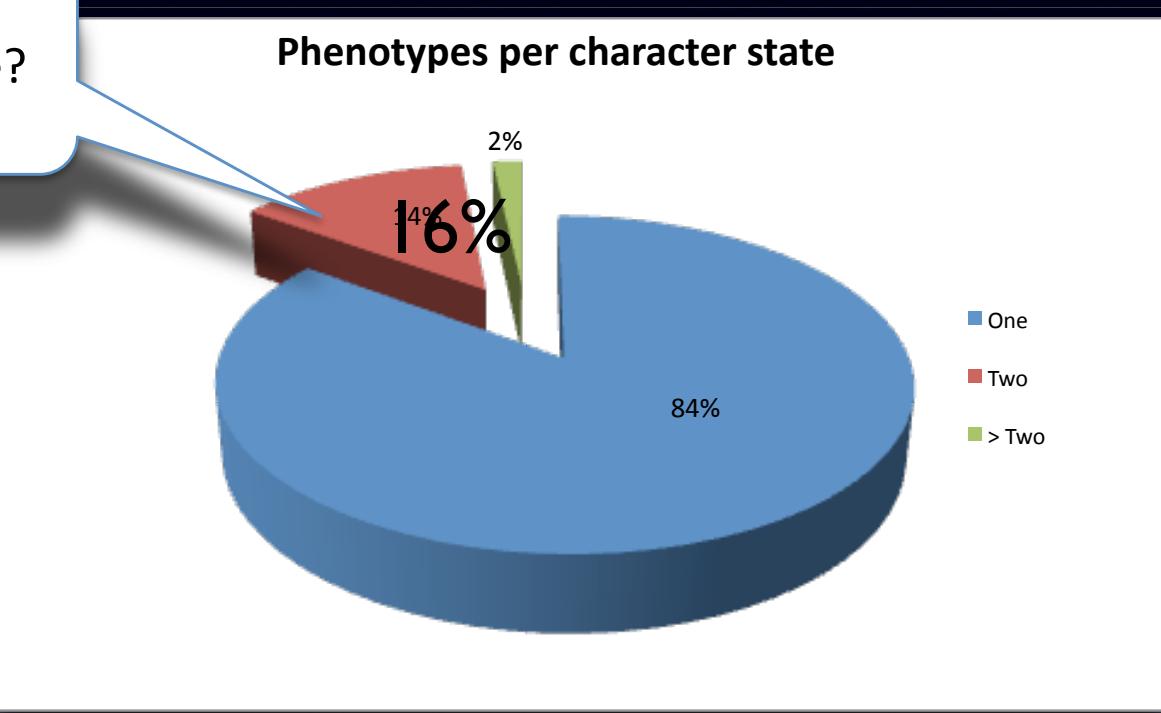
E: **primordial ligament**; Q: **size, narrow**

E1: **primordial ligament**; Q: **attached to**; E2: **maxilla ascending process**

# 1 character state: >1 phenotype

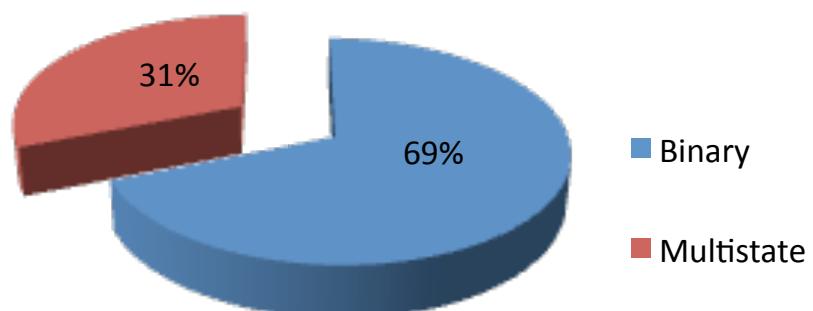
16%

Significance?

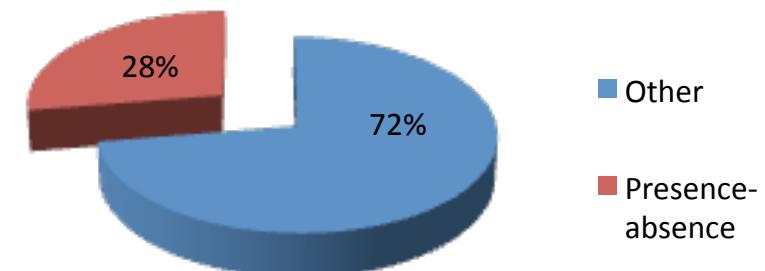


**Binary characters 69%;  
Presence-absence characters: 28%**

**Binary vs. Multistate**



**Presence-absence vs. Other**



# Relational vs. Non-relational phenotypes

Character 80 (Zanata & Vari, 2005). “Form and area of attachment of **primordial ligament**: (0) ligament relatively **narrow** and **attaching to** posteromedial portion of **ascending process of maxilla**.....

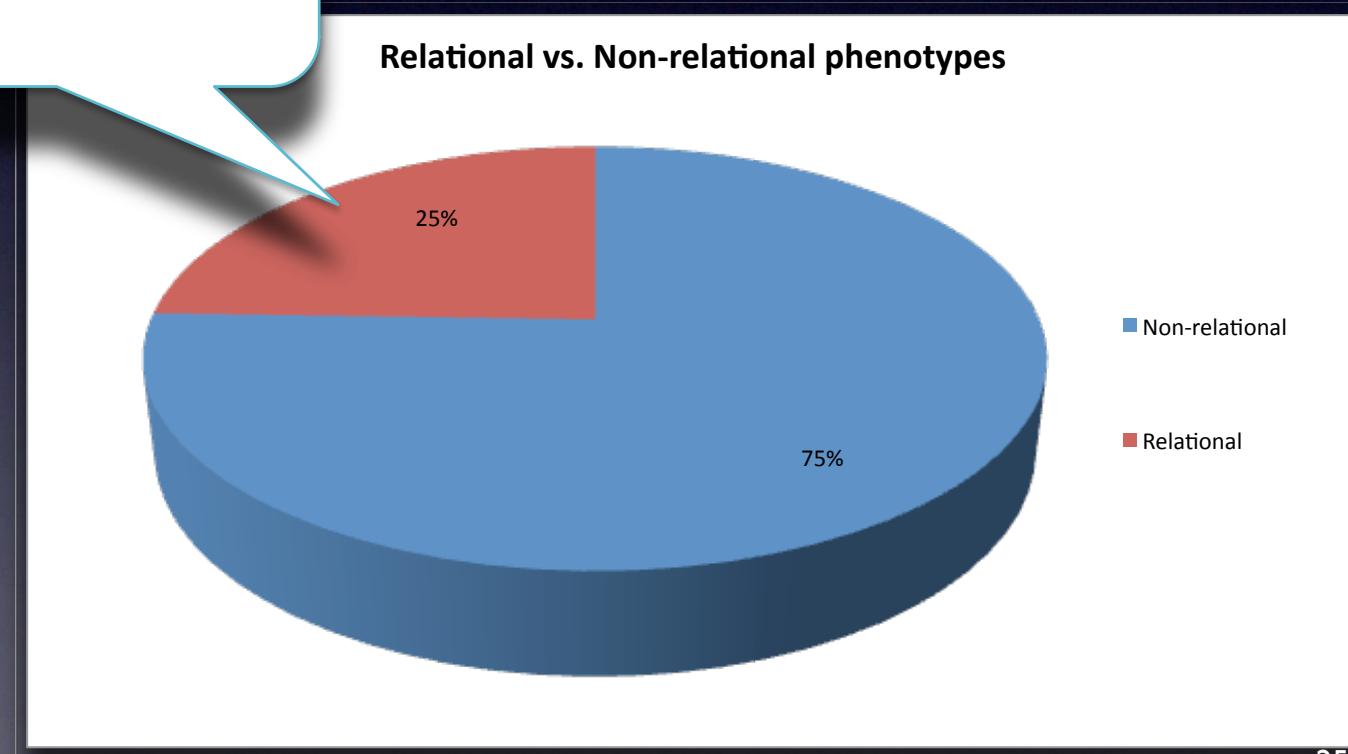
## 2 Phenotypes:

E: **primordial ligament**; Q: **size, narrow**

E1: **primordial ligament**; Q: **attached to**; E2: **maxilla ascending process**

# Relational phenotypes: 25%

*A attached\_to B*  
*A separated\_from B*



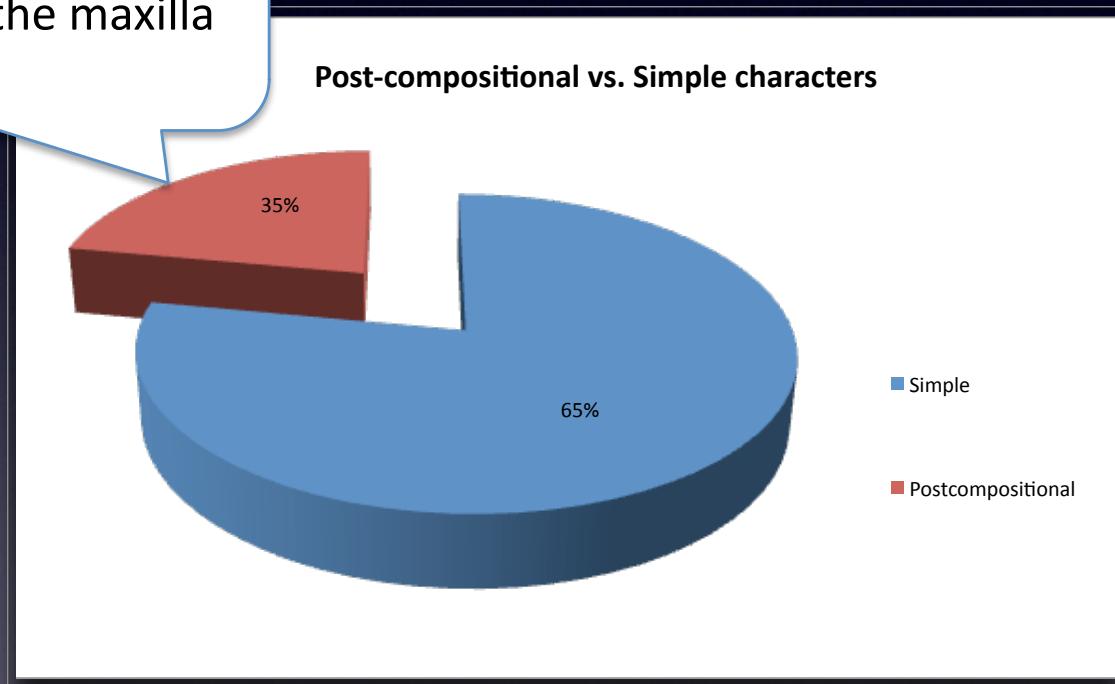
# Post-compositional characters: e.g., process *part\_of* maxilla

The screenshot shows the Phenoscape phenotype database interface. At the top, there is a logo with three green wavy lines and the word "PHENOSCAPE". Below the logo is a navigation bar with links for "Search", "Provide Feedback", and "About". A welcome message says "Welcome to the Phenoscape phenotype database for ostariophysan fishes." Below this, a section titled "Search the Phenoscape Knowledgebase" asks "Begin typing to choose a search term from the popup menu." A search input field contains the word "process". To the right of the input field, a list of search results is displayed, including terms like "ache", "ahctf1", "Amblyceps mangoi", "Ameiurus nebulosus", "Amphilus ateuensis", "anterior limb of parapophysis", "Ariopsis lelia", "Aspredinichthys thibeni", "Aspredo aspredo", "atp1a1", "Austroglanis barnardi", "autopalatine", "autopalatine-lateral ethmoid joint", "Bagre marinus", "Bagroides melapterus", "Bagrus bajad", "Batasio batasio", "cdh2", "cdx4", "Cetopsis coeculenta", "Chaca chaca", "Chrysipterus auratus", "Chrysichthys longipinnis", "Chrysichthys nigrodigitatus", "Clarotes laticeps", "coronomeckelian", "dentine tooth", "Diplomystes chilensis", "dorsal fin spine", "epibranchial 1 element", "epibranchial 2 element", "epibranchial 3 bone", "uncinate process", "epibranchial 5 cartilage", "ethmoid cartilage", "exti3", "eya1", "fgf18a", "frontal bone", "gill raker", "gpc4", "hdac1", "Hemibagrus nemurus", "hyomandibula", "ictalurus punctatus", "infraorbital 1", "kita", "lama1", "lamb1", "lamc1", "lateral ethmoid", "Loricaria cataphracta", "lrn6", "maxilla", "mesethmoid bone", "metapterygoid", "mib", "mycop", "Meckel's cartilage", "Mystus nigricans", "Nematogenys inermis", "Notarius flavus", "ntla", "opercle os suspensorium", "Pangasianodon hypophthalmus", "Parakysis verrucosus", "parapophysis", "pax2a", "posterior process of basipterygium", "pou5f1", "premaxilla", "Pseudolaguvia shawi", "Rhabdalestes septentrionalis", "Rhamdia quelen", "rib", "Rita rita", "scale", "Schilbe mystus", "sox10", "sox9a", "sp1", "smo", "tachysurus fulvidraco", "tdo", "tp53", "transcapular ligament", "trpm7", "yang12", "vomer", "wnt11", "wnt5b". Below the search results, the word "process" is displayed in large green letters, and the word "blocess" is partially visible below it. At the bottom of the page, there is a footer with various links and text.

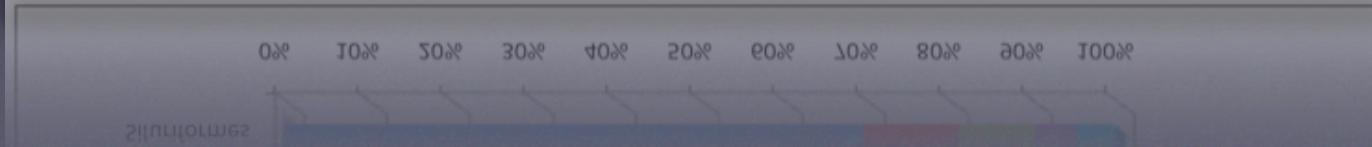
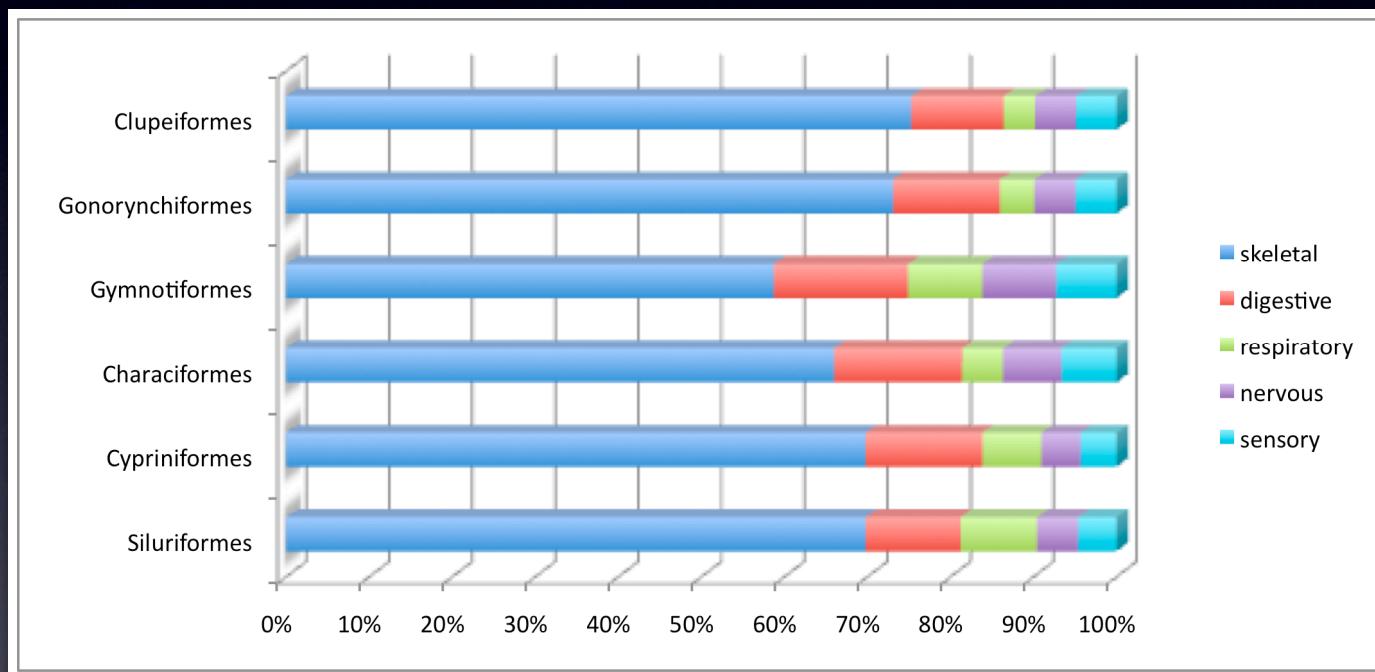
25 July 2008

# Doo-dads: 35%

e.g., anterior process of the lateral edge of the maxilla



# Distribution of all characters across anatomical systems in taxa



25 July 2008

# Distribution of skeletal characters in broad regions across taxa

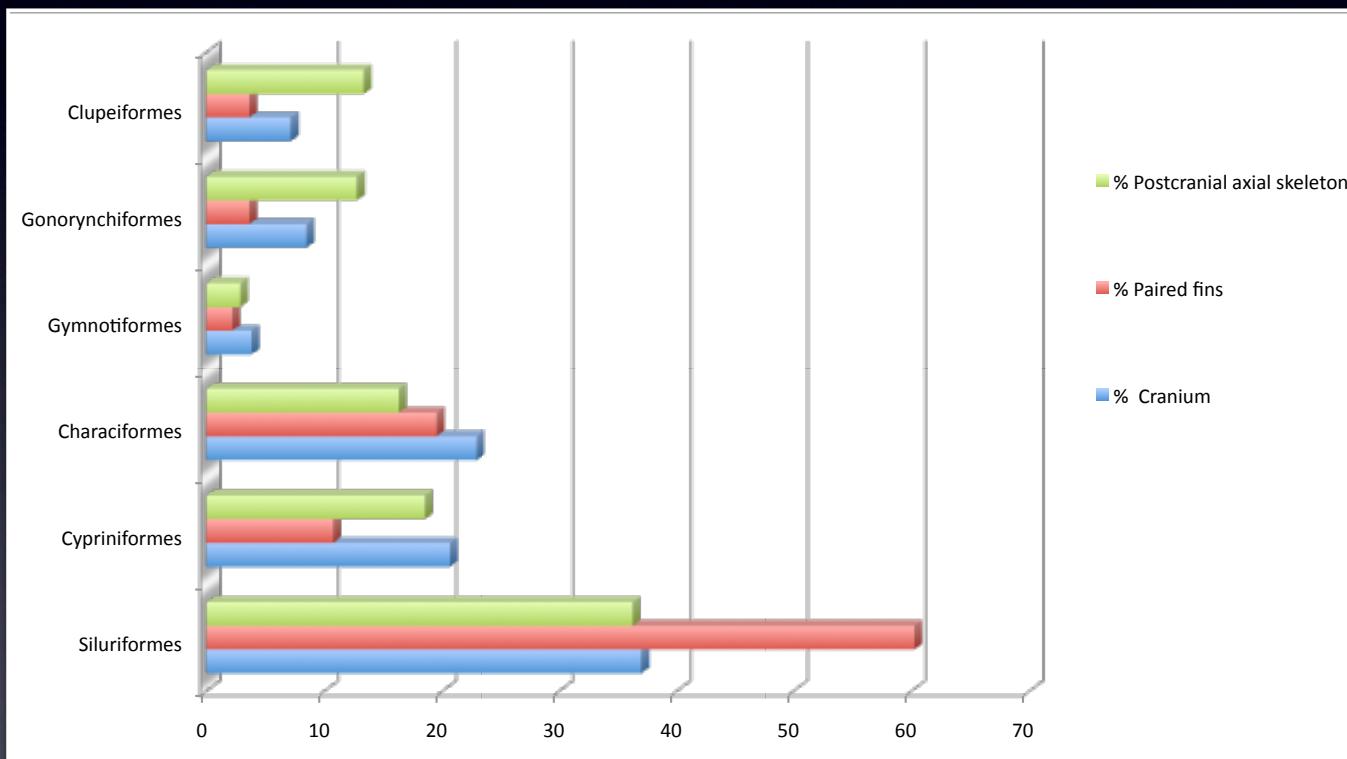
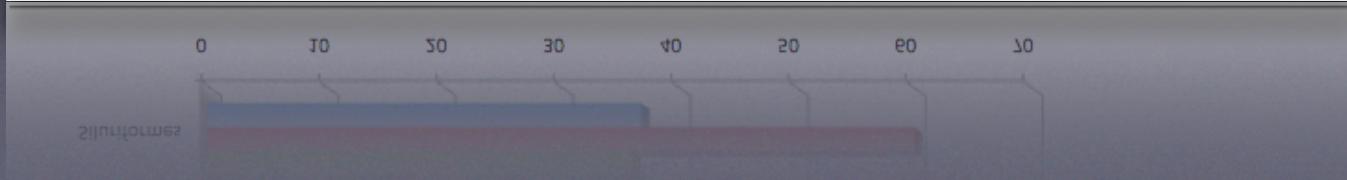


Image from Sabaj-Perez



25 July 2008

# Devo-Evo applications

# Search for zebrafish genes involved in fish scale development

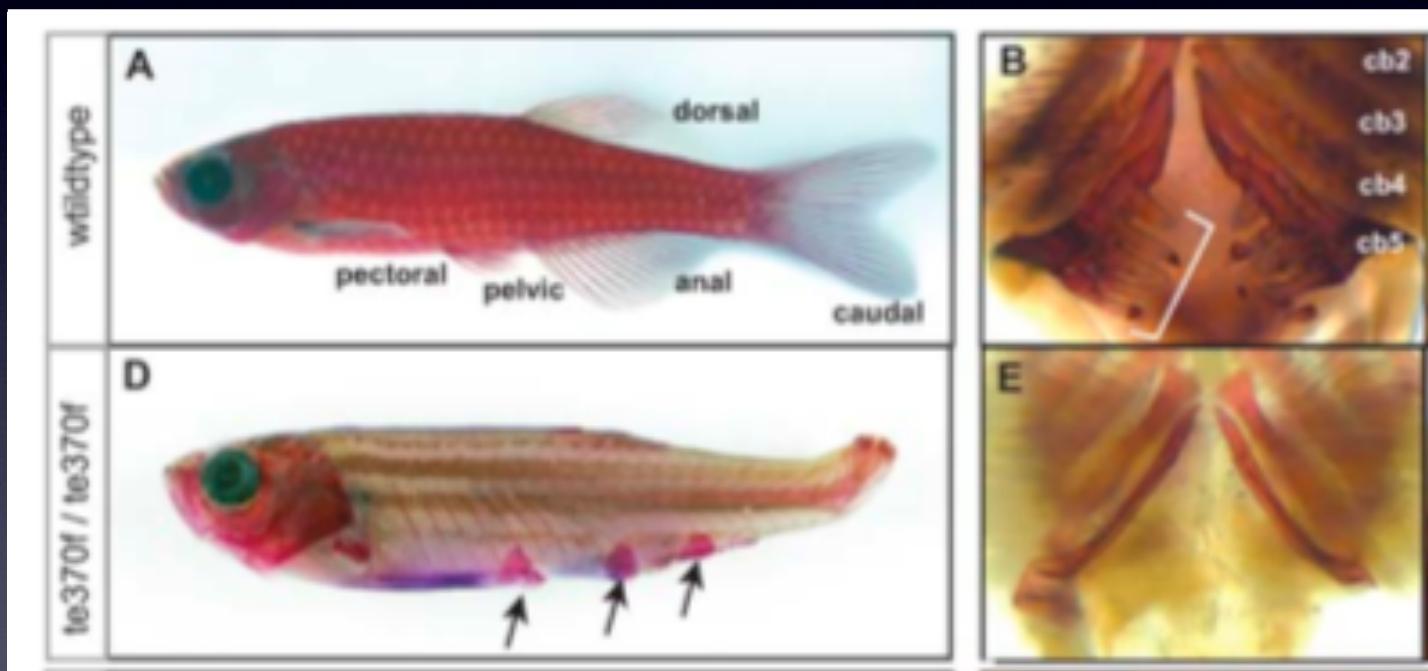
## Phenotypes

<u>Anatomy</u>	<u>Quality</u>	<u>Zebrafish Genes</u>	<u>Taxa</u>
<a href="#">scale</a>	<a href="#">relational spatial quality</a> : angular placement, relational spatial quality towards process of parietal bone, relational spatial quality towards supraoccipital crest	<a href="#">3</a>	<a href="#">19</a>
<a href="#">scale</a>	<a href="#">texture</a> : texture of	<a href="#">0</a>	<a href="#">91</a>
<a href="#">scale</a>	<a href="#">structure</a>	<a href="#">0</a>	<a href="#">95</a>
<a href="#">scale</a>	<a href="#">position</a> : inverted, spatial pattern	<a href="#">10</a>	<a href="#">79</a>
<a href="#">scale</a>	<a href="#">shape</a> : round	<a href="#">17</a>	<a href="#">372</a>
<a href="#">scale</a>	<a href="#">quality</a> : discontinuous, malformed, occurrence quality	<a href="#">6</a>	<a href="#">0</a>
<a href="#">scale</a>	<a href="#">count</a> : absent, count of 0, count of 0-3, count of 4-5, count of 6 and 3 more	<a href="#">5</a>	<a href="#">74</a>
<a href="#">scale</a>	<a href="#">size</a> : decreased size	<a href="#">5</a>	<a href="#">215</a>

Find 46 genes of interest

25 July 2008

# Zebrafish *finless* mutants: scale loss *eda* gene



Harris et al., 2007

# Search for fishes that lack scales:

## All phenotypes for scale count

- [-] Teleostei (645).....circuli absent, circuli count, circuli present, circuli of posterior surface of scale absent, circuli of posterior surface of scale present, lateral line scale absent, lateral line scale count, lateral line scale count, 29-32, lateral line scale count, 31-58, lateral line scale count, 35-41, lateral line scale count, 42-53, lateral line scale count, 55-65, lateral line scale count, mean: 125 (range: 87-139), lateral line scale count, mean: 142 (range: 130-168), lateral line scale count, mean: 82 (range: 78-115), lateral line scale count, ≤64, lateral line scale count, ≥33, lateral line scale count, ≥76, lateral line scale present, process of scale count, process of scale of dorsal side of caudal peduncle absent, process of scale of dorsal side of caudal peduncle present, process of scale of ventral side of caudal peduncle absent, process of scale of ventral side of caudal peduncle present, scale absent, scale count, 0, scale count, 0-3, scale count, 0-4, scale count, 3-4, scale count, 4-5, scale count, 6, scale of caudal fin absent, scale of caudal fin present, scale of caudal peduncle count, scale of caudal peduncle count, >17, scale of dorsal fin count, 30-39, scale of dorsal fin count, 40-49, scale of dorsal fin count, 50-56, scale on antero-dorsal region of body absent, scale on body absent, scale on body present, scale on dorsal region of body absent, scale on dorsal region of trunk absent, scale on head absent, scale on head present, scale on post-vent region count, 1-2, scale on post-vent region count, 4-5, scale on posterior region of head absent, scale on posterior region of head present, scale on postero-lateral region of body absent, scale on postero-lateral region of body present
- [-] Euteleostei (16).....circuli absent, circuli of posterior surface of scale absent, lateral line scale count, 31-58, lateral line scale count, mean: 125 (range: 87-139), lateral line scale count, mean: 142 (range: 130-168), lateral line scale count, mean: 82 (range: 78-115), lateral line scale present, scale absent, scale count, 0, scale count, 0-3, scale count, 0-4, scale count, 3-4, scale count, 4-5, scale count, 6
- [+] Protacanthopterygii (8).....circuli absent, lateral line scale count, 31-58, lateral line scale count, mean: 125 (range: 87-139), lateral line scale count, mean: 142 (range: 130-168), lateral line scale count, mean: 82 (range: 78-115), lateral line scale present
- [-] Neoteleostei (8).....circuli of posterior surface of scale absent, scale absent, scale count, 0, scale count, 0-3, scale count, 0-4, scale count, 3-4, scale count, 4-5, scale count, 6
- [-] Percomorpha (8).....circuli of posterior surface of scale absent, scale absent, scale count, 0, scale count, 0-3, scale count, 0-4, scale count, 3-4, scale count, 4-5, scale count, 6
  - [+] Perciformes (1).....circuli of posterior surface of scale absent
  - [-] Smegmamorpha (7).....scale absent, scale count, 0, scale count, 0-3, scale count, 0-4, scale count, 3-4, scale count, 4-5, scale count, 6
    - [-] Gasterosteiformes (7).....scale absent, scale count, 0, scale count, 0-3, scale count, 0-4, scale count, 3-4, scale count, 4-5, scale count, 6
      - [+] Hypoptychidae (1).....scale count, 6
      - [+] Aulorhynchidae (1).....scale count, 6
      - [-] Gasterosteidae (5).....scale absent, scale count, 0, scale count, 0-3, scale count, 0-4, scale count, 3-4, scale count, 4-5
        - [+] *Pungitius* (1).....scale count, 0-3
        - [+] *Culaea* (1).....scale count, 3-4
        - [+] *Apeltes* (1).....scale absent, scale count, 0
        - [+] *Spinachia* (1).....scale count, 4-5
        - [+] *Gasterosteus* (1).....scale count, 0-4

Find sticklebacks, catfishes, etc.

# Is *eda* involved in scale loss in catfishes or sticklebacks?



Copyright © Jean Ricardo Simões Vitule, All Rights Reserved

*Ictalurus punctatus*

?

Research needed

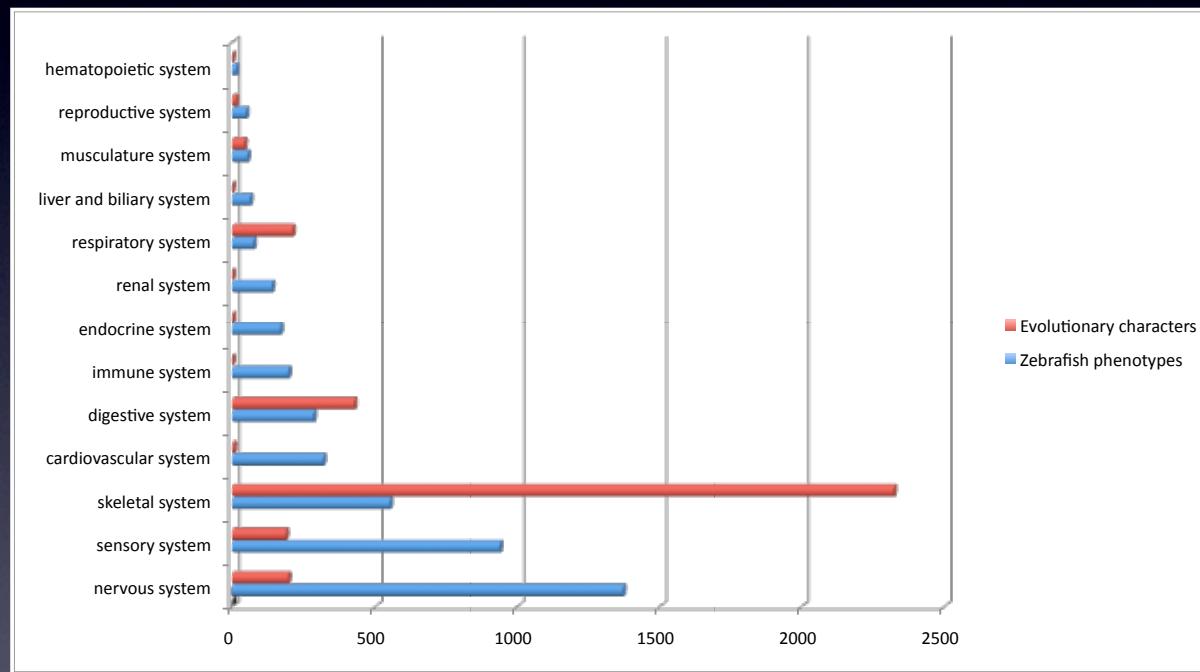
# Is *eda* involved in scale loss in sticklebacks?



*Gasterosteus aculeatus*

Yes -- *eda* is associated with scale loss in sticklebacks (Colosimo et al., 2005)

# Potential for identifying genetic basis of evolutionary characters



4,217 zebrafish phenotypes; 3,405 evolutionary characters



# Summary & Conclusions

- Developed taxonomy and anatomy ontologies
- Curated character data using ontologies
- Developed database and interface to integrate searching genes and anatomy
- Makes data accessible for broad group of researchers and creates opportunities for new and synthetic research



# Acknowledgements

National Science Foundation ( BDI-0641025)  
National Evolutionary Synthesis Center

## Contributors to Teleost Ontologies

Gloria Arratia  
Stan Blum  
Miles Coburn  
Kevin Conway  
Wasila Dahdul  
Mário de Pinna  
Jeff Engeman  
Bill Eschmeyer  
Terry Grande  
Melissa Haendel  
Brian Hall  
Eric Hilton  
John Lundberg  
Richard Mayden  
Mark Sabaj Pérez  
Brian Sidlauskas  
Richard Vari  
Jacqueline Webb  
Edward Wiley

**Curators:**  
Miles Coburn  
Jeff Engeman  
Terry Grande  
Eric Hilton  
John Lundberg  
Paula Mabee  
Richard Mayden  
Mark Sabaj

## Phenoscape Workshop Participants & Contributors

Arhat Abzhanov  
Michael Ashburner  
Judith Blake  
Stan Blum  
Quentin Cronk  
Mário de Pinna  
Andy Deans  
George Gkoutos  
Melissa Haendel  
Hopi Hoekstra  
Hans Hofmann  
Elizabeth Jockusch  
Elizabeth Kellogg  
Chuck Kimmel  
Suzanna Lewis  
Anne Maglia  
Austin Mast  
Chris Mungall  
Martin Ramirez  
Sue Rhee  
Martin Ringwald  
Nelson Rios  
Mark Sabaj Pérez  
Eric Segerdell  
Brian Sidlauskas  
Barry Smith  
David Stern  
Peter Vize  
Gunter Wagner  
Nicole Washington  
Edward Wiley

# More....

- Phenoscape poster Sunday evening
- Phenoscape hands-on demonstrations, this afternoon, Exec Towers, Salon II
- [phenoscape.org](http://phenoscape.org)
- [kb.phenoscape.org](http://kb.phenoscape.org)