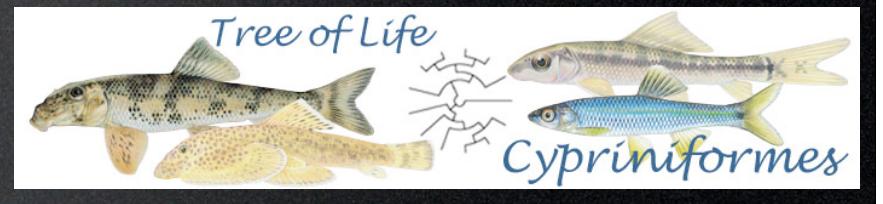
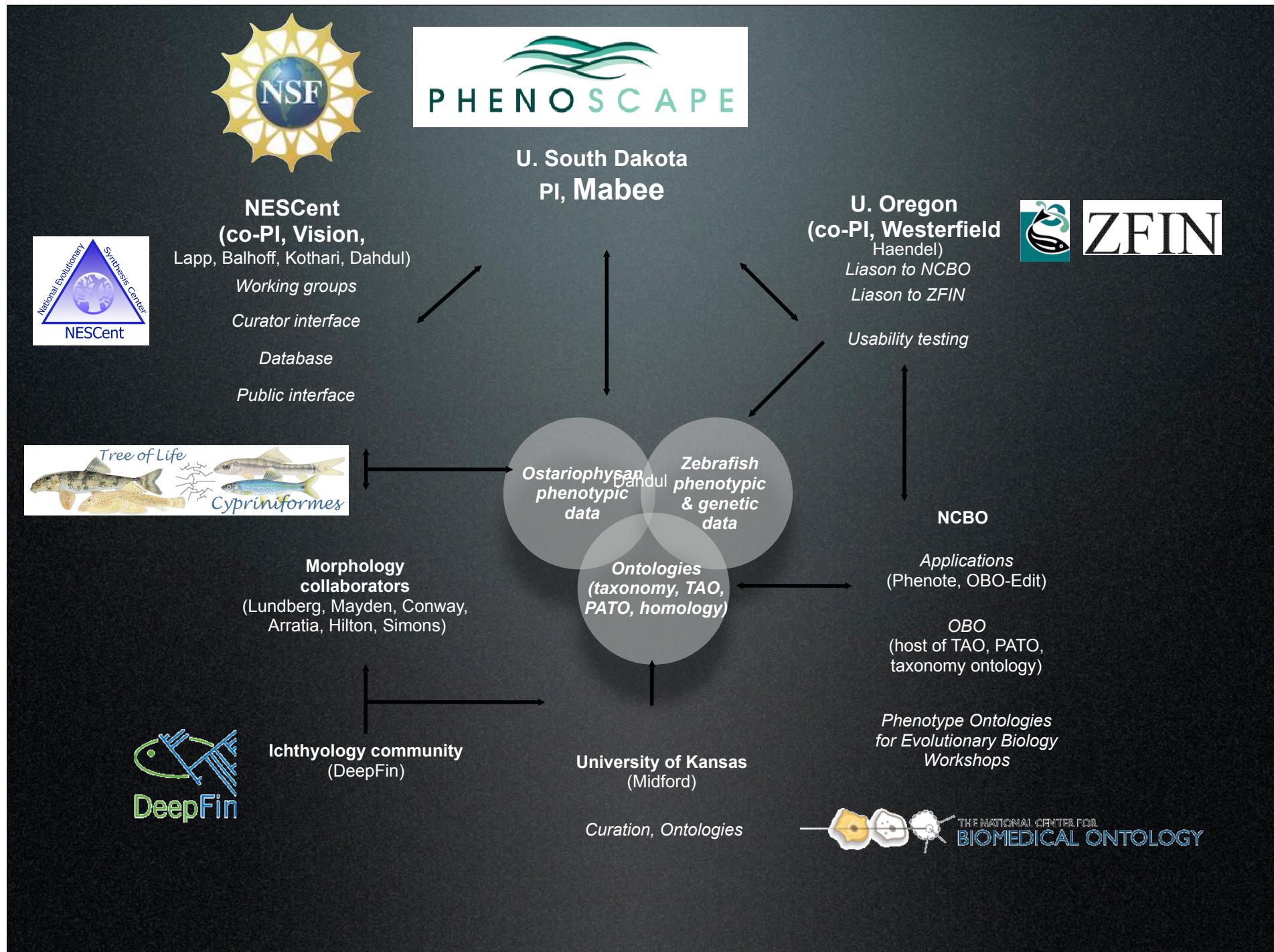


Phenoscape: Bridging systematics and genetics with ontologies

Paula Mabee
University of South Dakota







Goal: Connect these data using ontologies

Model
organisms
(biomed)



Comparative
evolutionary
biology

- Phenotypes
- Genetics

zebrafish

- Morphology
- Phylogeny

ostariophysan fishes

Ontologies can connect human disease to candidate genes

Animal models

Mutant gene
↓
Mutant or Missing Protein
↓
Mutant phenotype (disease model)

Humans

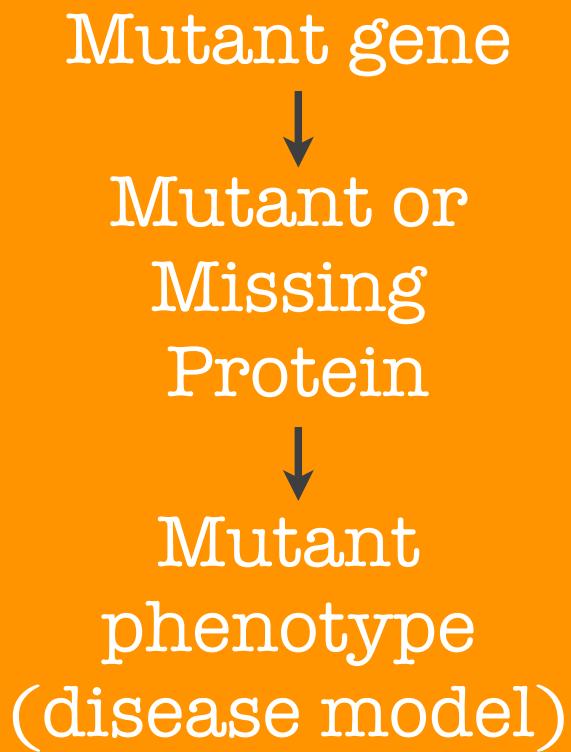
Mutant gene
↓
Mutant or Missing Protein
↓
Mutant phenotype (disease)

ONTOLOGIES

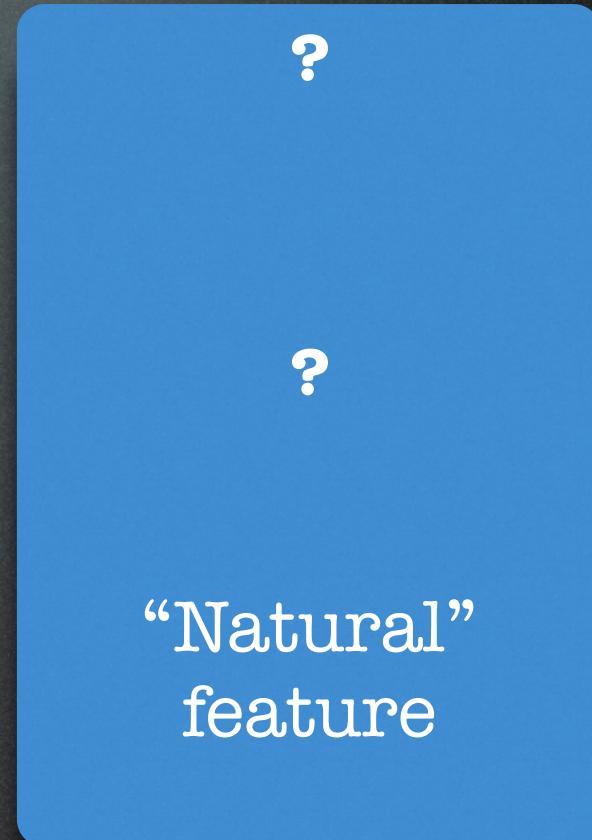


Ontologies can connect species variation to candidate genes

Animal models

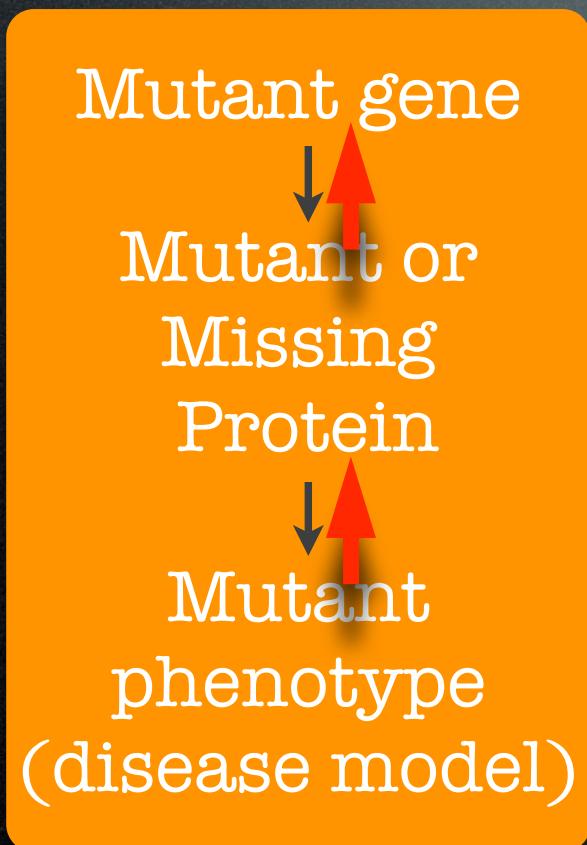


Evolutionary Species



Ontologies can connect species variation to candidate genes

Animal models



Evolutionary Species

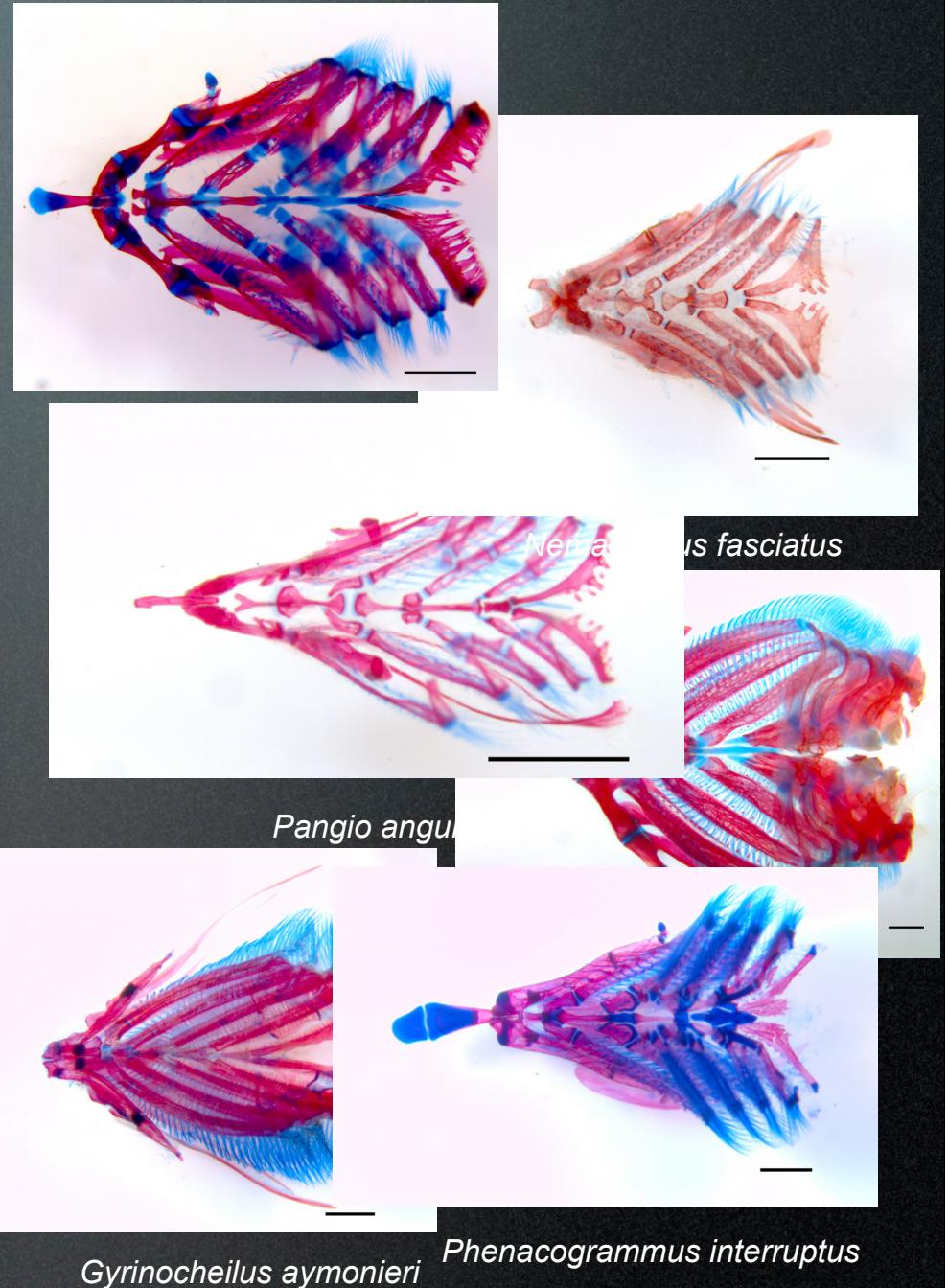
Candidate gene

Natural phenotype

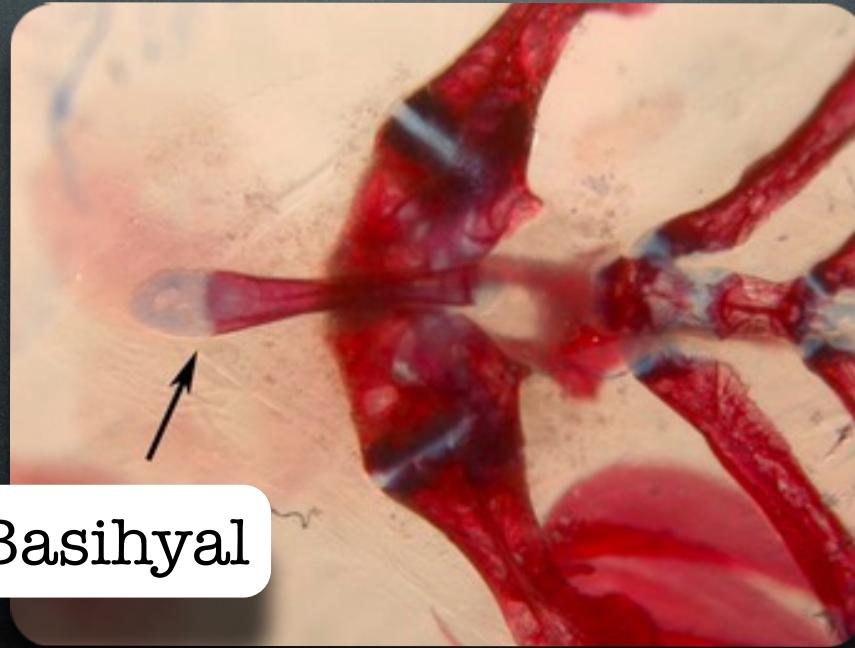


Phenoscape goals:

- identify taxa in which specific features vary
- identify candidate genes for evolutionary changes
- discover similar phenotypes and similar OTUs (phenotypic BLAST)
- correlation matrices of traits & genes



E.g. Basihyal lost in catfishes

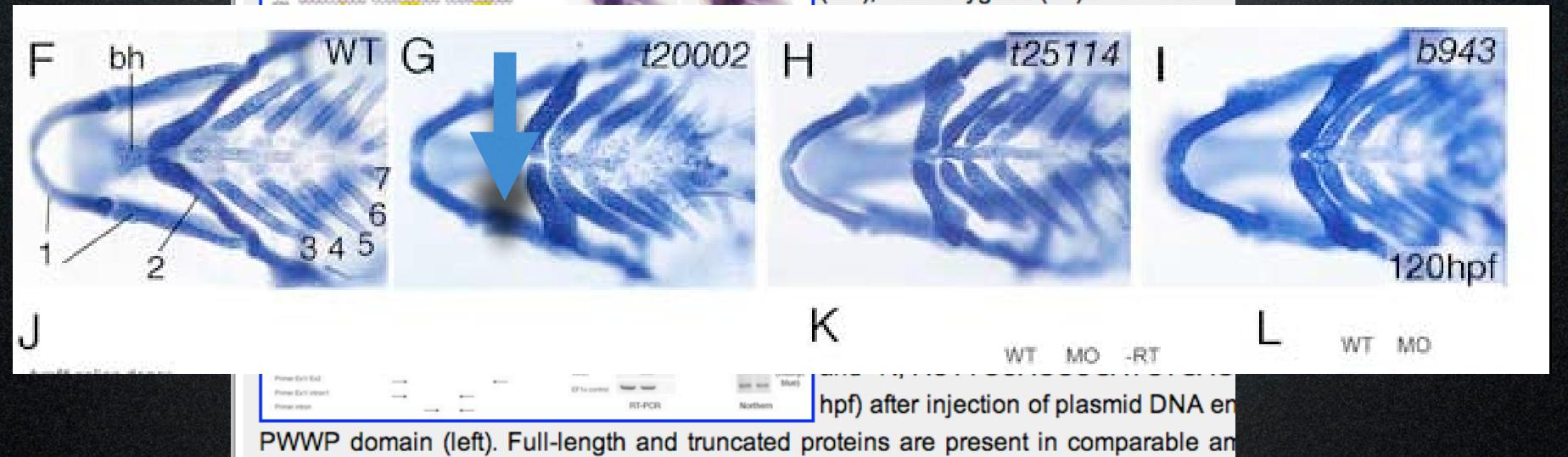


Basihyal primitively present

Sample query: What are all the possible genes underlying evolutionary loss in the basihyal?

Results: GENES ANATOMY IMAGES

PHENOTYPE affecting basihyal Mutant and Transgenic Lines			
Genotype (Background)	Affected Genes	Phenotype	Figures
brpf1^{b943/b943}	brpf1	absent, hypoplastic	2 figures from 1 publication
brpf1^{t20002/t20002}	brpf1	present	4 figures from 1 publication
brpf1^{t25114/t25114}	brpf1	absent, hypoplastic	2 figures from 1 publication
sox9a^{tw37/+}	sox9a	aplastic	1 figure from 1 publication
sox9a^{tw37/tw37}	sox9a	aplastic	1 figure from 1 publication



Basihyal loss in zebrafish due to mutation in *brpf1*

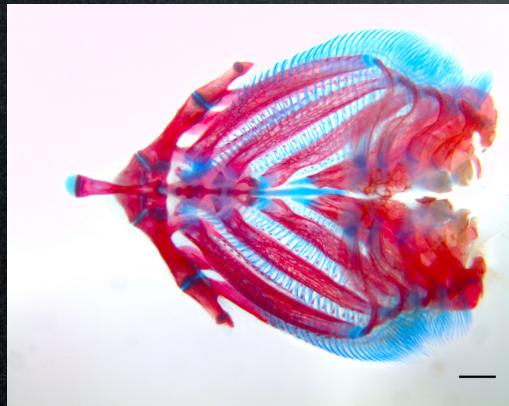
ZFIN ID: ZDB-FIG-080604-30

[Laue et al., 2008](#) - The multidomain protein Brpf1 binds histones and is required for Hox gene expression and segmental identity.
Development 135(11):1935-1946 - [Full text @ Development](#)

Requirements:

- **Develop ontologies**
- Develop appropriate syntax
- Curate phenotypes (characters)

Traditional comparative study: survey morphological variation



Cyprinus carpio



Pangio anguillaris



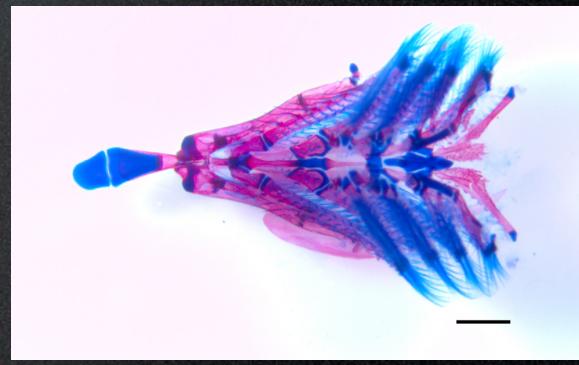
Nemacheilus fasciatus



Catostomus commersoni



Gyrinocheilus aymonieri



Phenacogrammus interruptus

Morphological data in free text format

- not a computable format
- cannot be compared among species
- cannot be linked to underlying genetics
- cannot be reasoned across

APPENDIX 1: CHARACTERS USED FOR PHYLOGENETIC ANALYSIS

Unless otherwise indicated, terminology follows Ronquist (1995a) and Ronquist and Nordlander (1989). Transformation series hypotheses are given for multi-state characters. Following each character is the character's consistency index and retention index on the preferred tree (Fig. 4). Observed character states are given in Table 3.

General Body Sculpture

1. Microsculpture on vertex, lateral surface of pronotum and mesoscutum: (0) absent, surface not dull (Figs. 9A–9D and 10A–10C); (1) present, linear, making the surface dull (not illustrated): (CI = 1.00, RI = 1.00, goodness of fit (G-fit) = 10).

Head

2. Shape of head in anterior view: (0) rounded, approximately as high as broad (Figs. 8A, 8B, and 9C); (1) elongate, higher than broad (Figs. 8C, 8D, 9A, and 9B): (CI = 0.25, RI = 0.82, G-fit = 5).

3. Relative position of eye: (0) close to ocelli, ratio of distance between compound eye and posterior mandibular articulation to distance between posterior ocellus and compound eye ≥ 1.18 (Figs. 8B and 8C); (1) removed from ocelli, ratio ≤ 1.13 (not illustrated); (CI = 0.20, RI = 0.50, G-fit = 4.3).

4. Size of ocelli: (0) small, ratio of maximum diameter of a lateral ocellus to shortest distance between lateral ocelli 0.22–0.40 (not illustrated); (1) large, ratio 0.44–0.65 (Figs. 8B and 8D); (CI = 0.11, RI = 0.62, G-fit = 2.7).

(Figs. 8B and 8D); (1) long (not illustrated); (CI = 0.20, RI = 0.33, G-fit = 4.3).

7. Shape of compound eyes in dorsal view: (0) rounded, distinctly protruding from the surface of the head, particularly anteriorly (Figs. 8B and 8D); (1) less rounded, not distinctly protruding from the surface of the head (not illustrated); (CI = 0.25, RI = 0.73, G-fit = 5).

8. Lateral frontal carina: (0) absent (Fig. 8D); (1) present (Fig. 8B, more easily seen in dorsal view); (CI = 0.50, RI = 0.50, G-fit = 7.5).

9. Hair punctures on lateral part of vertex: (0) indistinct or absent (Figs. 8B and 8D); (1) present, distinctly enlarged (not illustrated); (CI = 0.33, RI = 0.60, G-fit = 6).

10. Sculpture on posterior part of vertex (seen in dorsal view, not illustrated): (0) smooth or punctate, without linear component; (1) with parallel or slightly radiating, transverse strigae; (CI = 0.50, RI = 0.75, G-fit = 7.5).

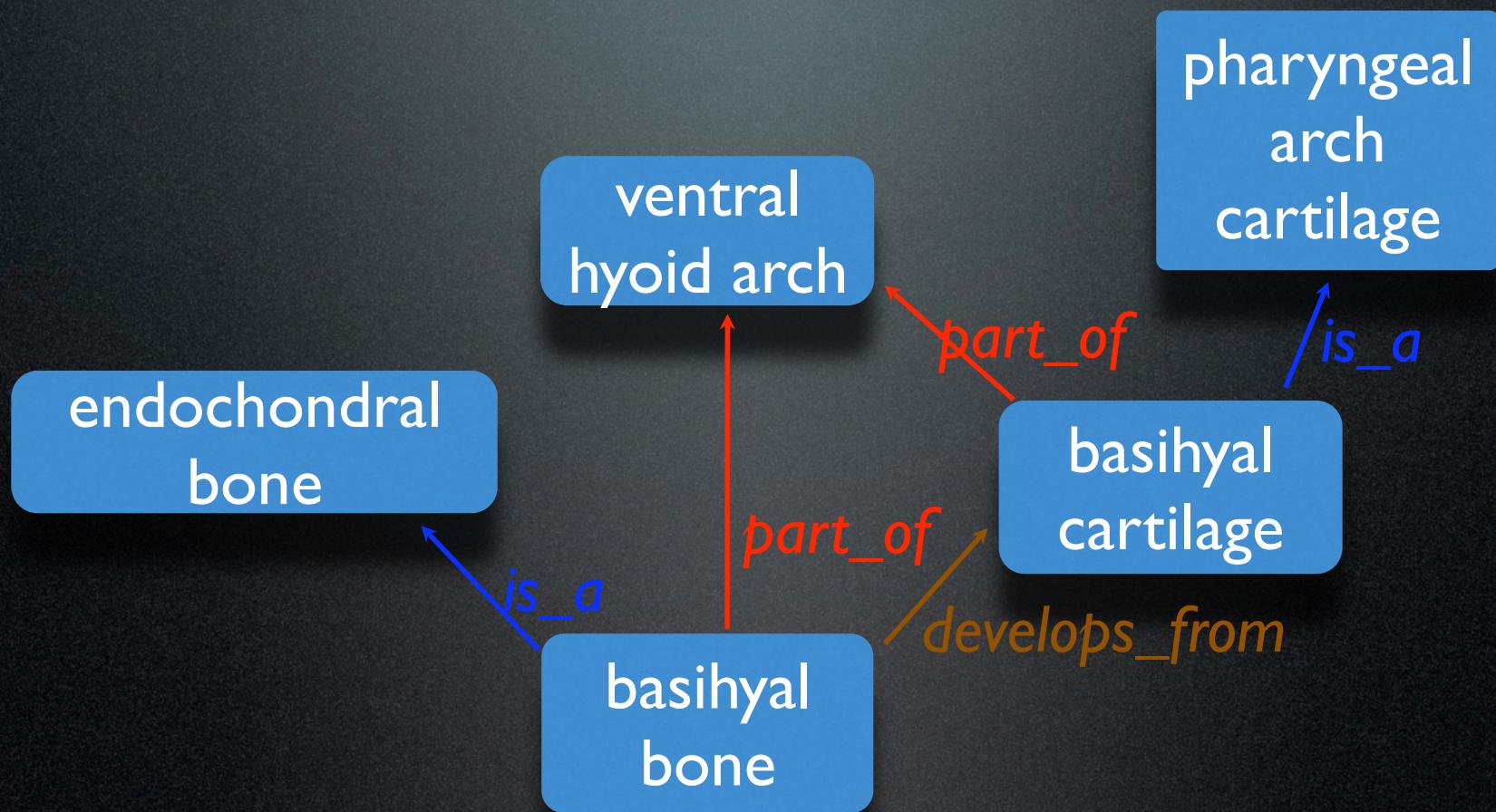
11. Relative position of antennal sockets: (0) close to ocelli; ratio of vertical distance between inner margin of antennal foramen and ventral margin of clypeus to vertical distance between anterior ocellus and antennal rim < 2.0 (not illustrated); (1) intermediate, ratio 2.25–4.1 (Figs. 8B and 8D); (2) far from ocelli, ratio > 4.4 (not illustrated). Ordered 012; (CI = 0.08, RI = 0.33, G-fit = 12).

12. Vertical carina adjacent to ventral margin of antennal socket: (0) absent (Fig. 8B); (1) present (Fig. 8D); (CI = 0.50, RI = 0.86, G-fit = 7.5).

13. Vertical delineations on lower face: (0) absent (Figs. 8B and 8D); (1) single carina or ledge (not illustrated); (2) several parallel or subparallel carinae (not illustrated). Unordered; (CI = 0.29, RI = 0.54, G-fit = 3.8).

14. (Subdivision of 13:1) Shape of single vertical delineation of lower face (not illustrated): (0) rounded divergent ledges running from antennal sockets to dor-

Teleost Anatomy Ontology



Phenoscape ontologies

New:

Teleost Anatomy
Ontology
(2233 terms; 387
skeletal)

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

Taxonomic
Rank
Ontology
(8->31 terms)

Zebrafish
Anatomy
Ontology
(2196 terms; 310
skeletal)

Existing:

Phenotype and
Trait Ontology
(1,075 terms)

Spatial
Ontology
(106 terms)

Evidence
Code
Ontology

Requirements:

- Develop ontologies
- **Develop appropriate syntax**
- Curate phenotypes (characters)

Requirements:

- Develop ontologies
- Develop appropriate syntax
- **Curate phenotypes (characters) from [prototype Otophysi] fishes**

Creating prototype by curating evolutionary data

Taxon	# Species	# Papers	# Characters
Cypriniformes (Mayden; Coburn)	3,268	70	1125
Siluriformes (Lundberg)	2,867	87	1200
Characiformes (Dahdul)	1,674	124	800
Gymnotiformes (Arratia)	134	2	200
Gonorynchiformes (Arratia)	37	80	75
Clupeiformes (Hilton)	364	60	380
TOTAL	8,344	423	3,780

63,080, 640 phenotypic data points

Summary

- Evolutionary morphology and genetic data can be linked using ontologies
- **Phenoscape**: prototype of a generalizable system for making this connection
- Crosses disciplines, facilitates new discoveries through data mining

Future goal: Connect many new data types using ontologies

Model
organisms
(biomed)



Comparative
evolutionary
biology

- Phenotypes
- Genetics

- Morphology
- Phylogeny
- Ecology
- Locality

Training opportunity

- 10 week internships
 - travel + stipend
 - typically summer (May-August)
- Mabee lab or National Evolutionary Synthesis Center
- informatics; morphological data; ontologies