

Phenoscape:

Extending model organism
ontologies for devo-evo studies
of evolutionary phenotypes

Paula Mabee

University of South Dakota



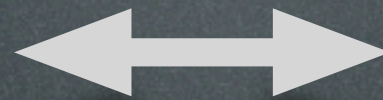


Goal: Connect these data using ontologies

Model
organisms
(biomed)

- Phenotypes
- Genetics

zebrafish



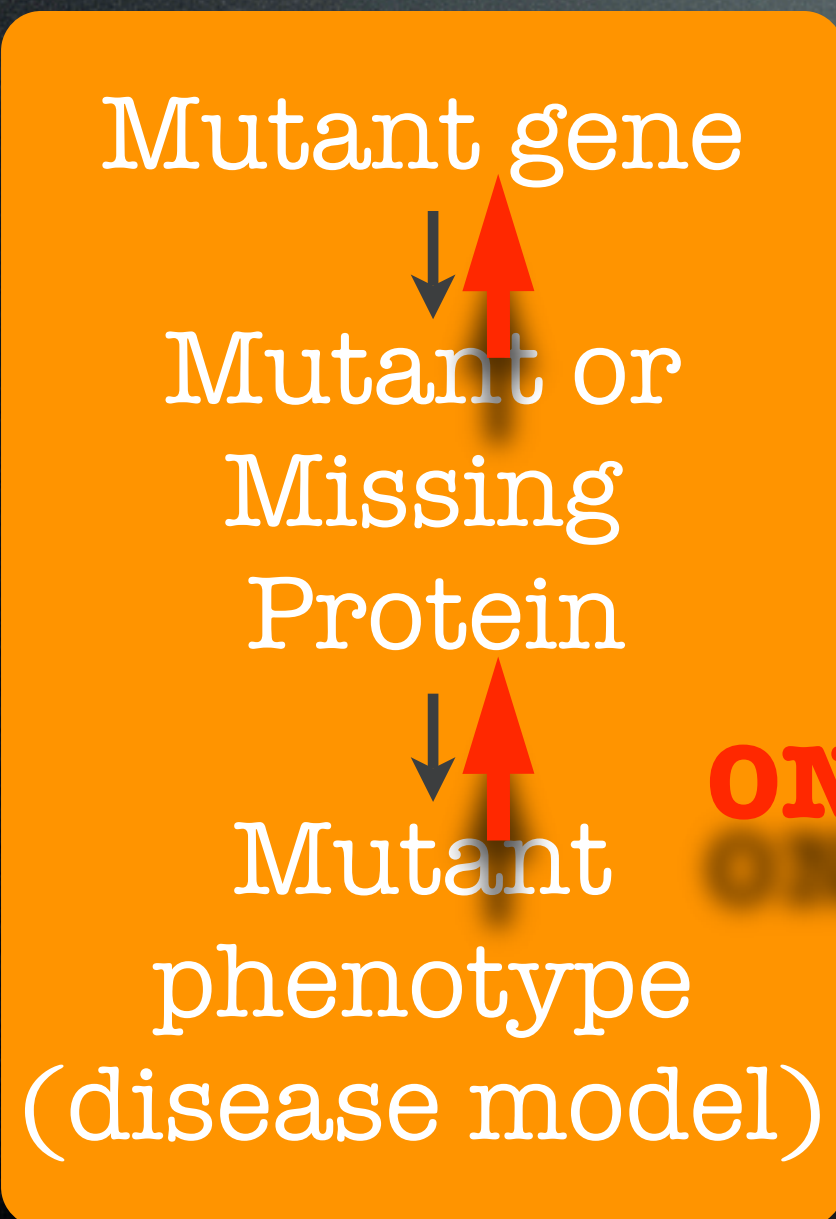
Comparative
Evolutionary
Biology

- Morphology
- Phylogeny
- Homology
- Taxa

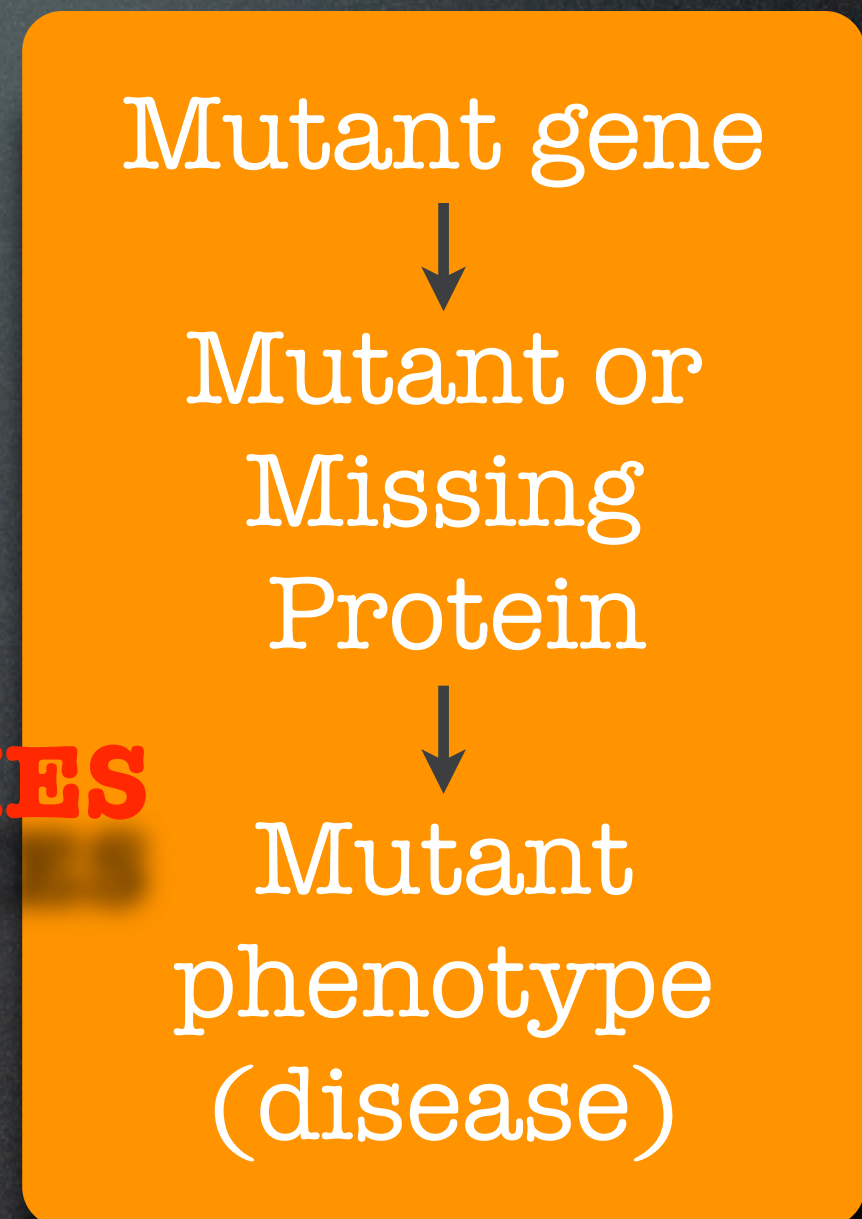
ostariophysan fishes

Ontologies can connect human disease to candidate genes

Animal models



Humans



ONTOLOGIES



Zebrafish ↔ Human

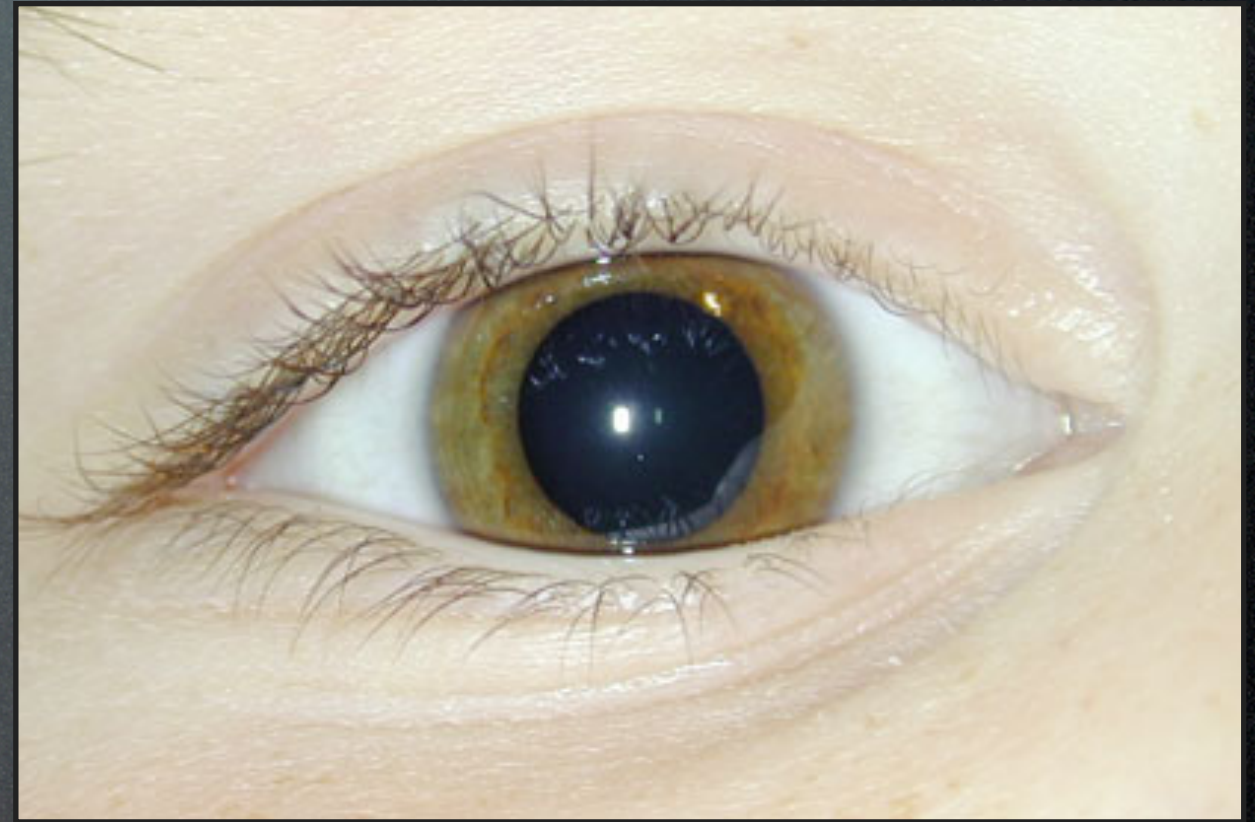
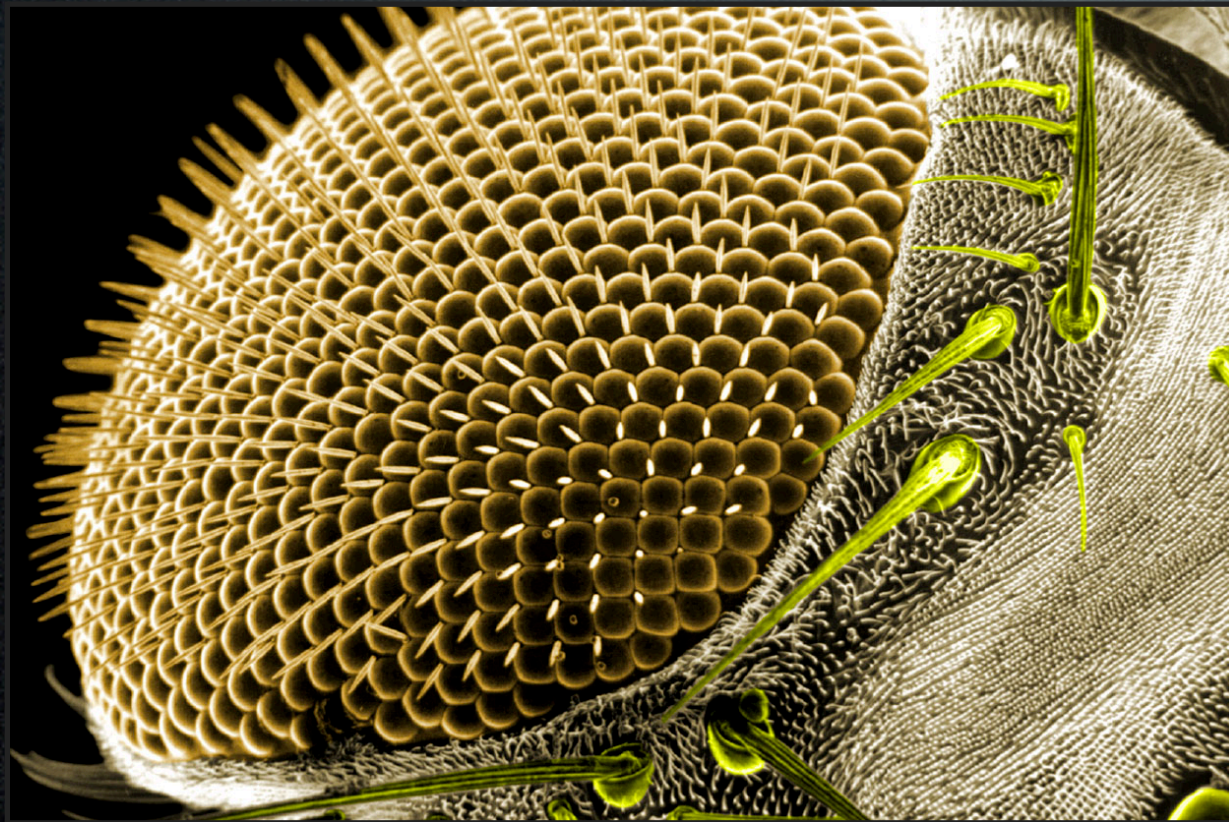
Conservation of gene sequence & function



slc24a5 involved in pigmentation similarity between fish and humans (Lamason et al., 2005)

Fly ↔ Human

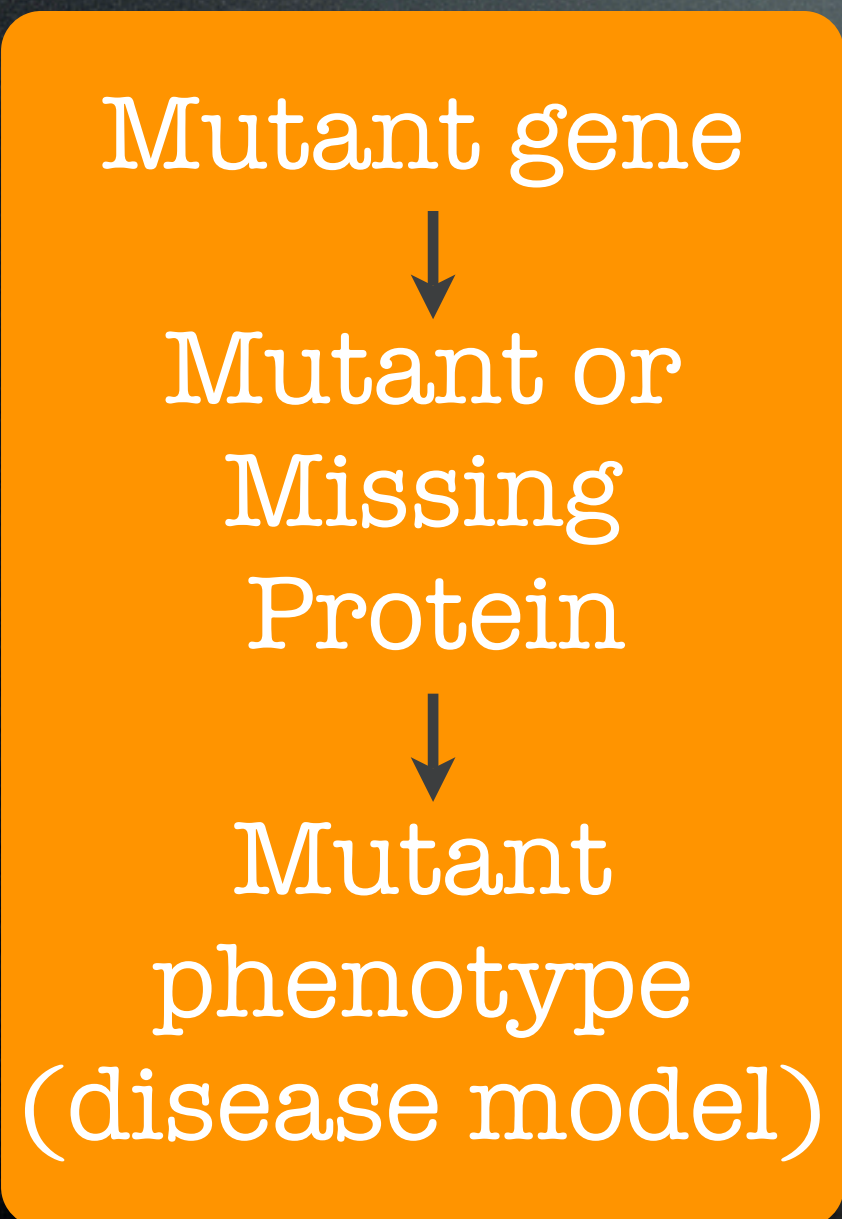
Conservation of developmental genes & toolkit



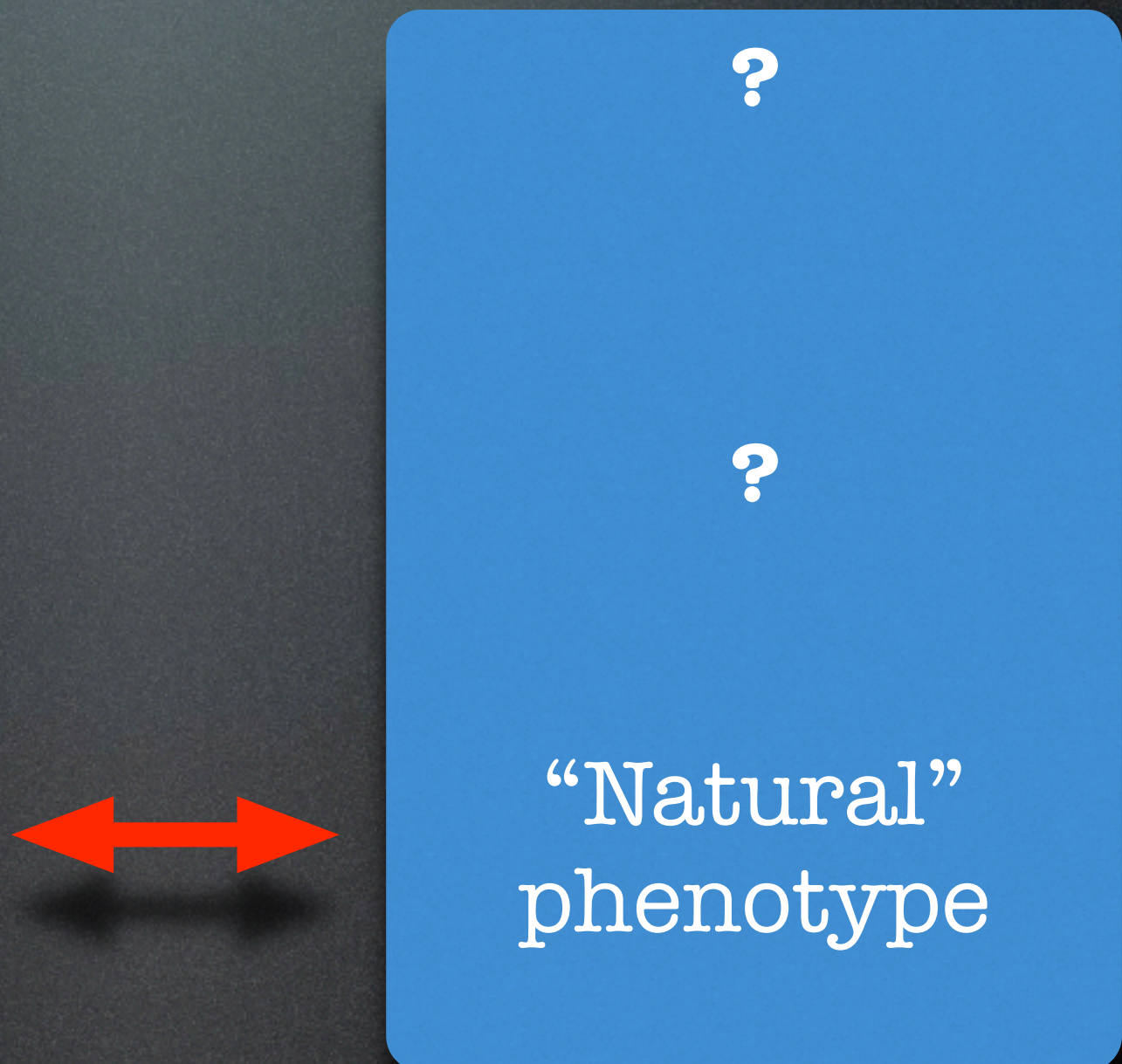
Pax6/eyeless gene directs development of fly and human eyes

Problem: Lack of genetic data for most species

Animal models



Evolutionary Species

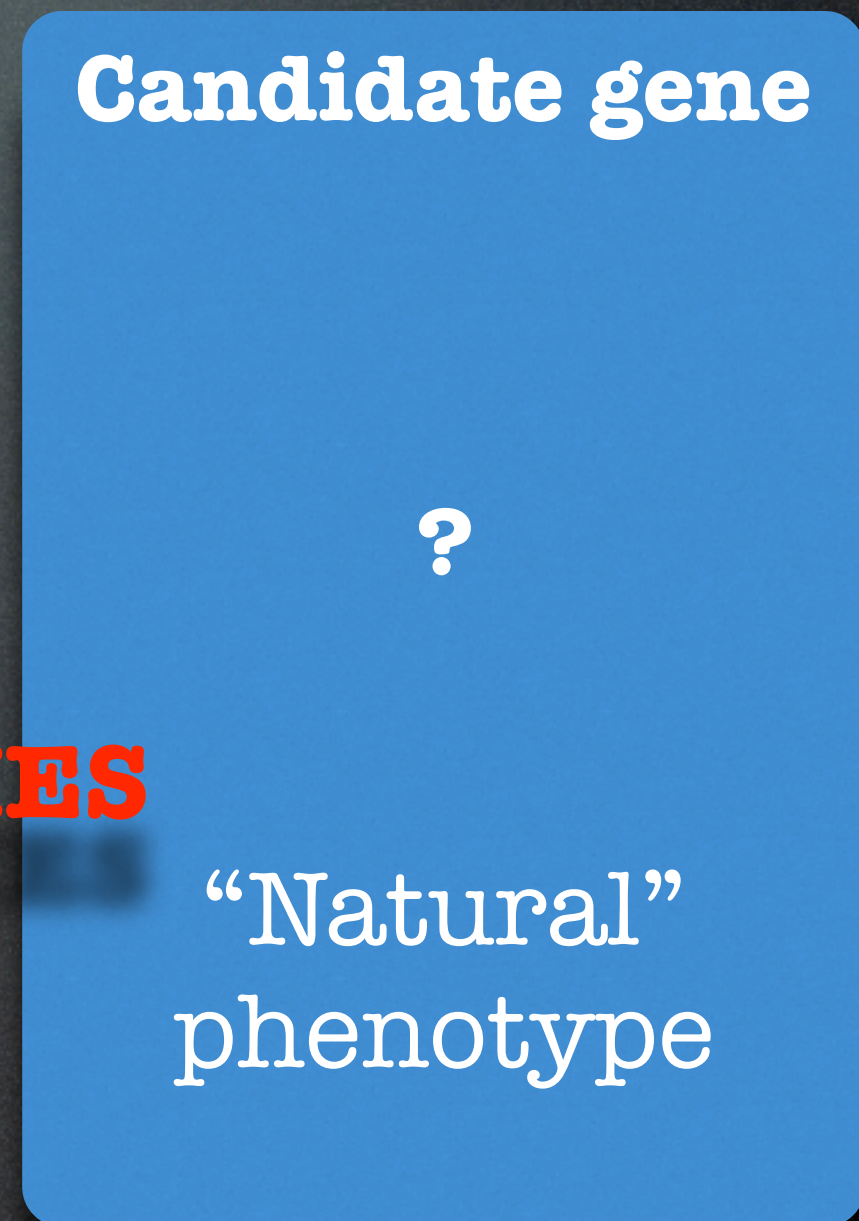


Ontologies can connect species variation to candidate genes

Animal models



Evolutionary Species



ONTOLOGIES

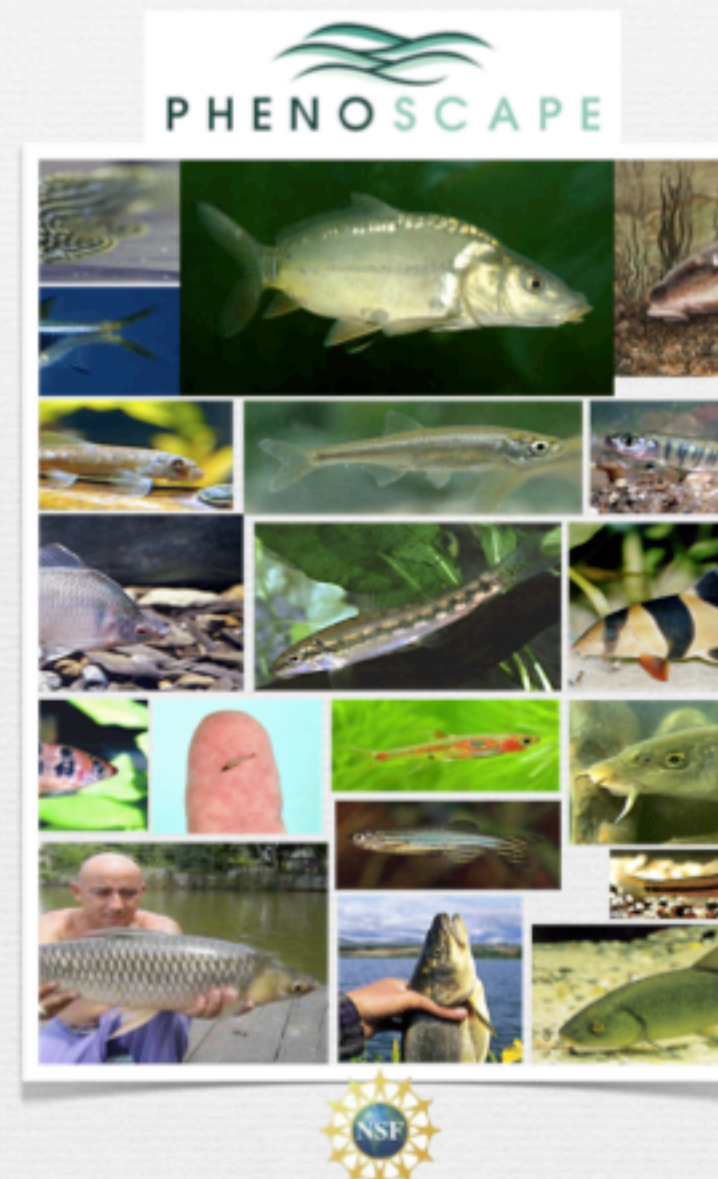




- Goals:
 - identify candidate genes for evolutionary phenotypes
 - mapping and identification of taxa in which specific features vary
 - discover similar phenotypes and similar OTUs (phenotypic BLAST)
 - correlation matrices of traits & genes

Phenoscape.org

- **History:** Communication between zebrafish model organism community and Cypriniformes Tree of Life group through NESCent workshops (Mabee-Westerfield)
- **Goal:** Create curated, ontology-based evolutionary phenotype database that maps to genetic databases
- **Generalizable system:** Prototype with ostariophysan fishes





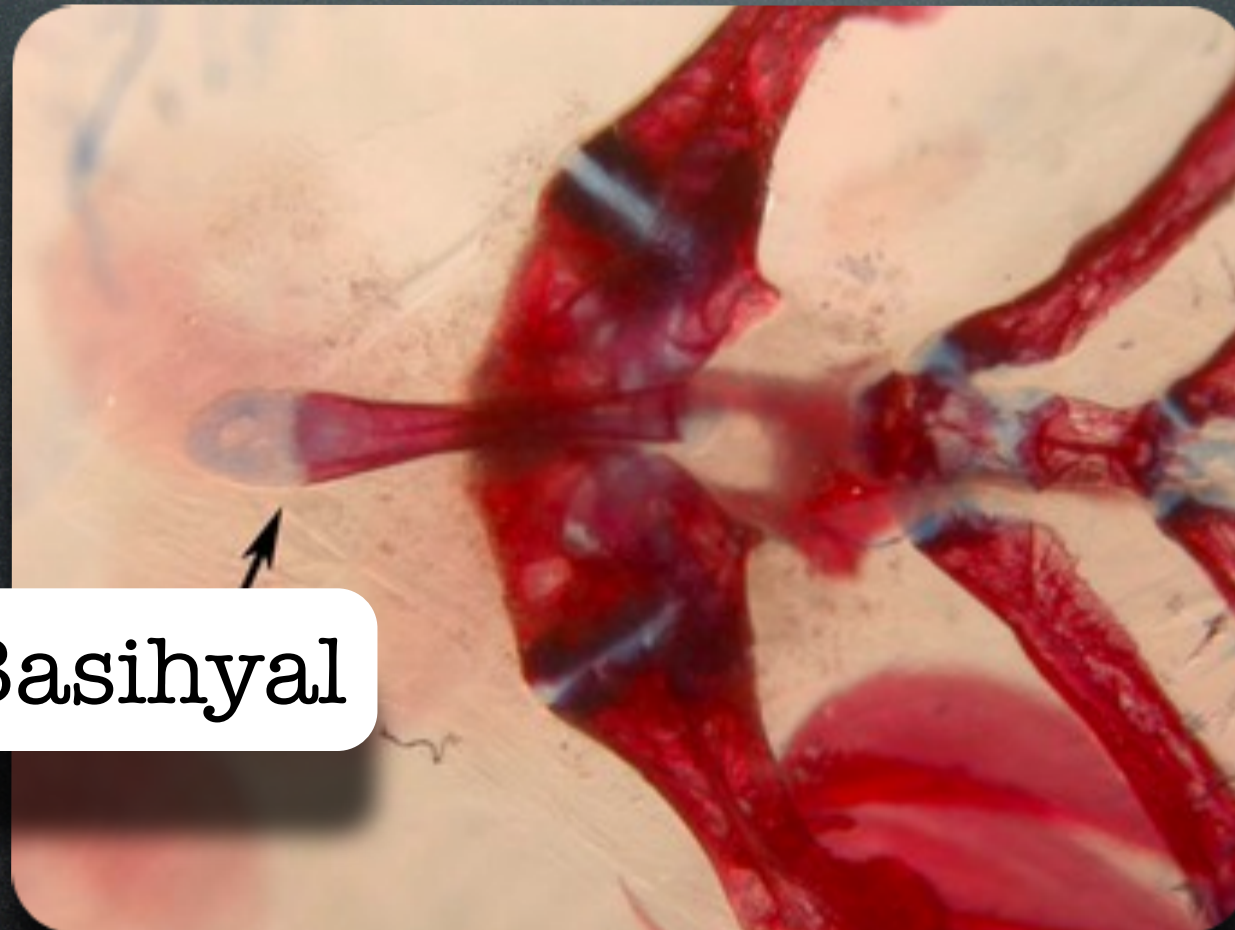
Ostariophysi



Ostariophysi: Diverse, speciose, freshwater fishes including zebrafish

Use case: Basihyal lost in catfishes

(Basihyal primitively present)



Basihyal

Sample query: What are all the possible genes underlying evolutionary loss of 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

PHENOTYPE:

Genotype(s): [brpf1^{b943/b943}](#) ▼, [brpf1^{t20002/t20002}](#) ▼, [brpf1^{t25114/t25114}](#) ▼

Observed In: [basihyal](#)

Stage Range : [Day 5](#)



Fig. S1 Analysis of *brpf1* mutations (+/+), heterozygous (+/-) and *t20002* mutant (-/-) embryos. (A) Sequencing profiles of *brpf1* in wild-type (+/+), heterozygous (+/-) and *t20002* mutant (-/-) embryos. The *t20002* mutation is indicated by an asterisk. (B) Sequencing profile of *brpf1* in the *t25114* mutant (-/-, lower panel). The *t25114* mutation is indicated by an asterisk. (C) Whole-mount in situ hybridization for *brpf1* in wild-type (WT) and *t25114* -/- embryos at 48 hpf. (D) Whole-mount in situ hybridization for *brpf1* in a *t20002* mutant embryo at 48 hpf. (E) Whole-mount in situ hybridization for *brpf1* in a *t25114* mutant embryo at 48 hpf. (F) Whole-mount in situ hybridization for *brpf1* in a *b943* mutant embryo at 120 hpf. (G) Whole-mount in situ hybridization for *brpf1* in a *t20002* mutant embryo at 120 hpf. (H) Whole-mount in situ hybridization for *brpf1* in a *t25114* mutant embryo at 120 hpf. (I) Whole-mount in situ hybridization for *brpf1* in a *b943* mutant embryo at 120 hpf. (J) Whole-mount in situ hybridization for *brpf1* in a wild-type (WT) embryo at 120 hpf. (K) Whole-mount in situ hybridization for *brpf1* in a wild-type (WT), mutant (MO), and reverse transcriptase (RT) control embryo at 120 hpf. (L) Whole-mount in situ hybridization for *brpf1* in a wild-type (WT) and mutant (MO) embryo at 120 hpf.

Basihyal loss in zebrafish due to mutation in *brpf1* gene

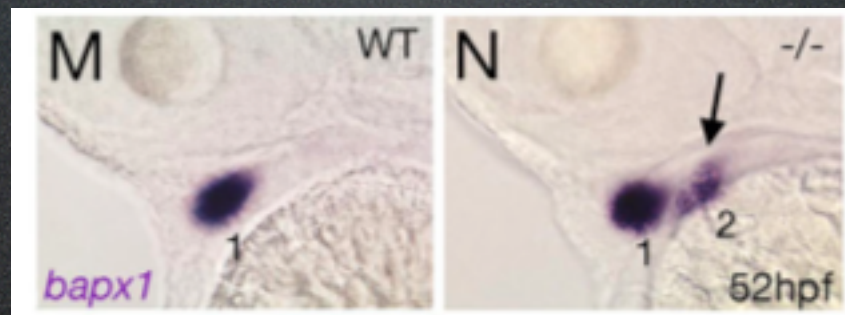
ZFIN ID: ZDB-FTG-080604-30

[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](#)

New hypothesis:

-Evolutionary loss of basihyal in catfishes due to changes in *brpf1* or Hox gene

-**Test** by looking at predicted gene expression in catfish



In the *brpf1* zebrafish mutant, *bapx1* is expressed in the 2nd pharyngeal arch

Requirements:

- Develop ontologies
- **Annotate evolutionary phenotypes**
 - using ontologies
 - using a syntax shared with model organism databases
- Develop curation software

Traditional comparative study: survey morphological variation



Cyprinus carpio



Pangio anguillaris



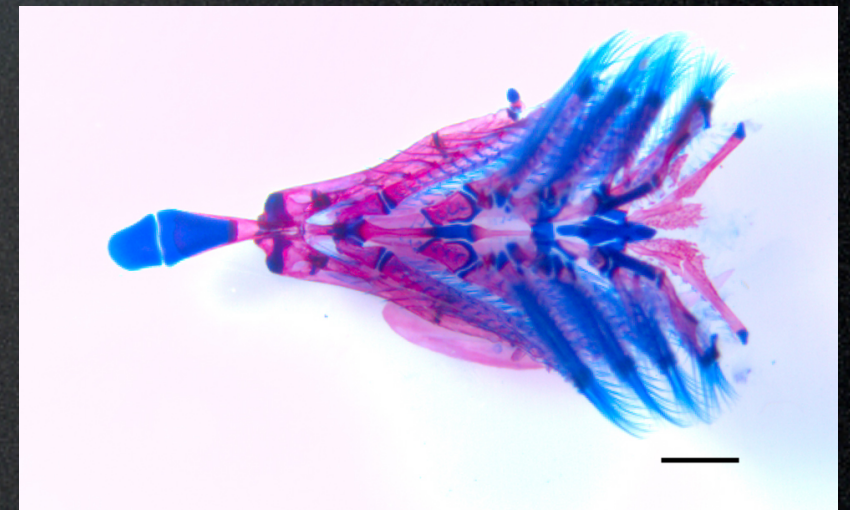
Nemacheilus fasciatus



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).

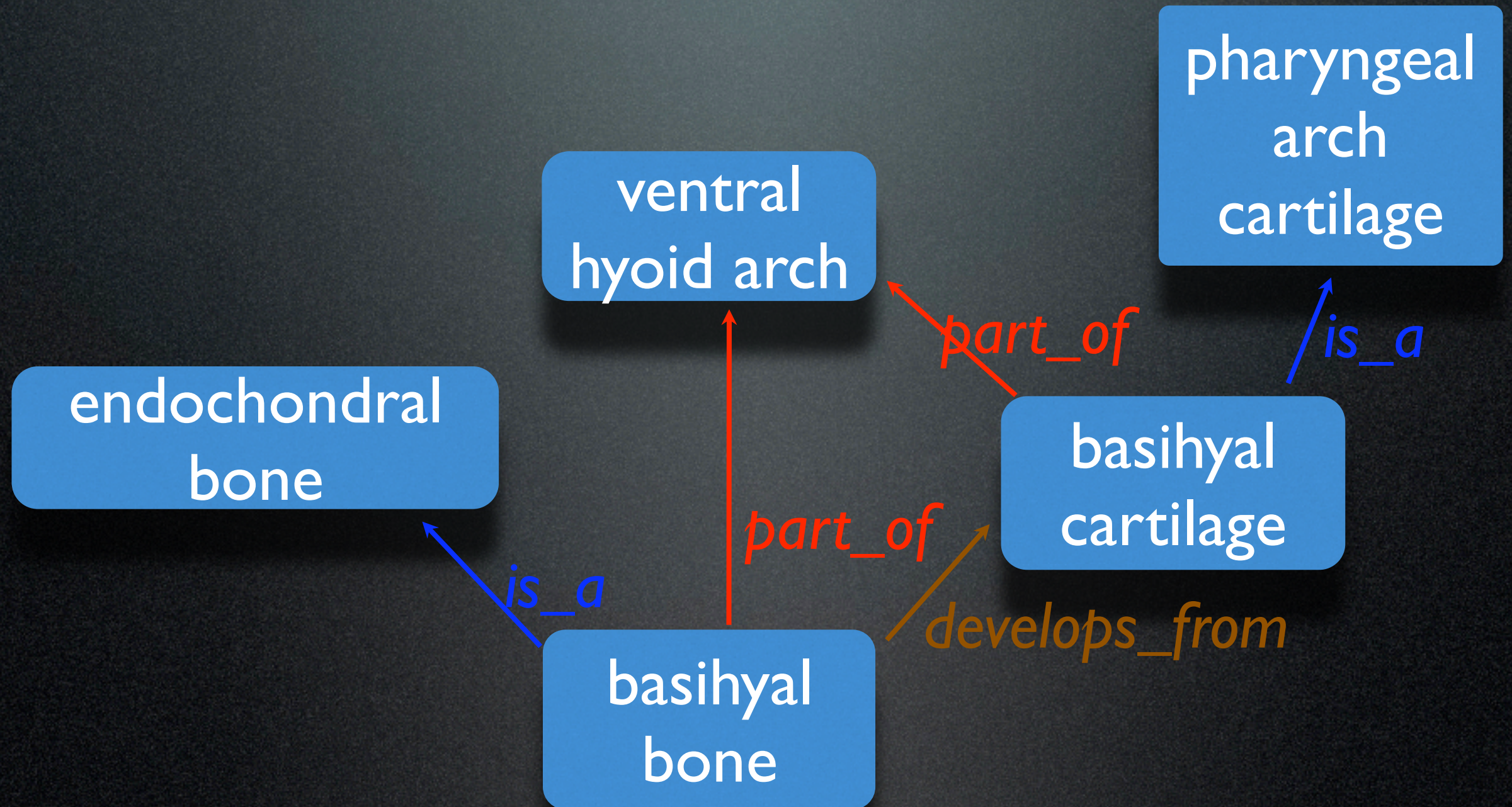
11. Relative position of antennal sockets: (0) close to ocelli; ratio of vertical distance between inner margin of antennal foramen and ventral margin of clypeus to vertical distance between anterior ocellus and antennal rim < 2.0 (not illustrated); (1) intermediate, ratio 2.25–4.1 (Figs. 8B and 8D); (2) far from ocelli, ratio > 4.4 (not illustrated). Ordered 012; (CI = 0.08, RI = 0.33, G-fit = 1.2).

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
(2371 terms; 395
skeletal)

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

Taxonomic
Rank
Ontology
(8->31 terms)

Existing:

Phenotype and
Trait Ontology
(1,075 terms)

Spatial
Ontology
(106 terms)

Evidence
Code
Ontology

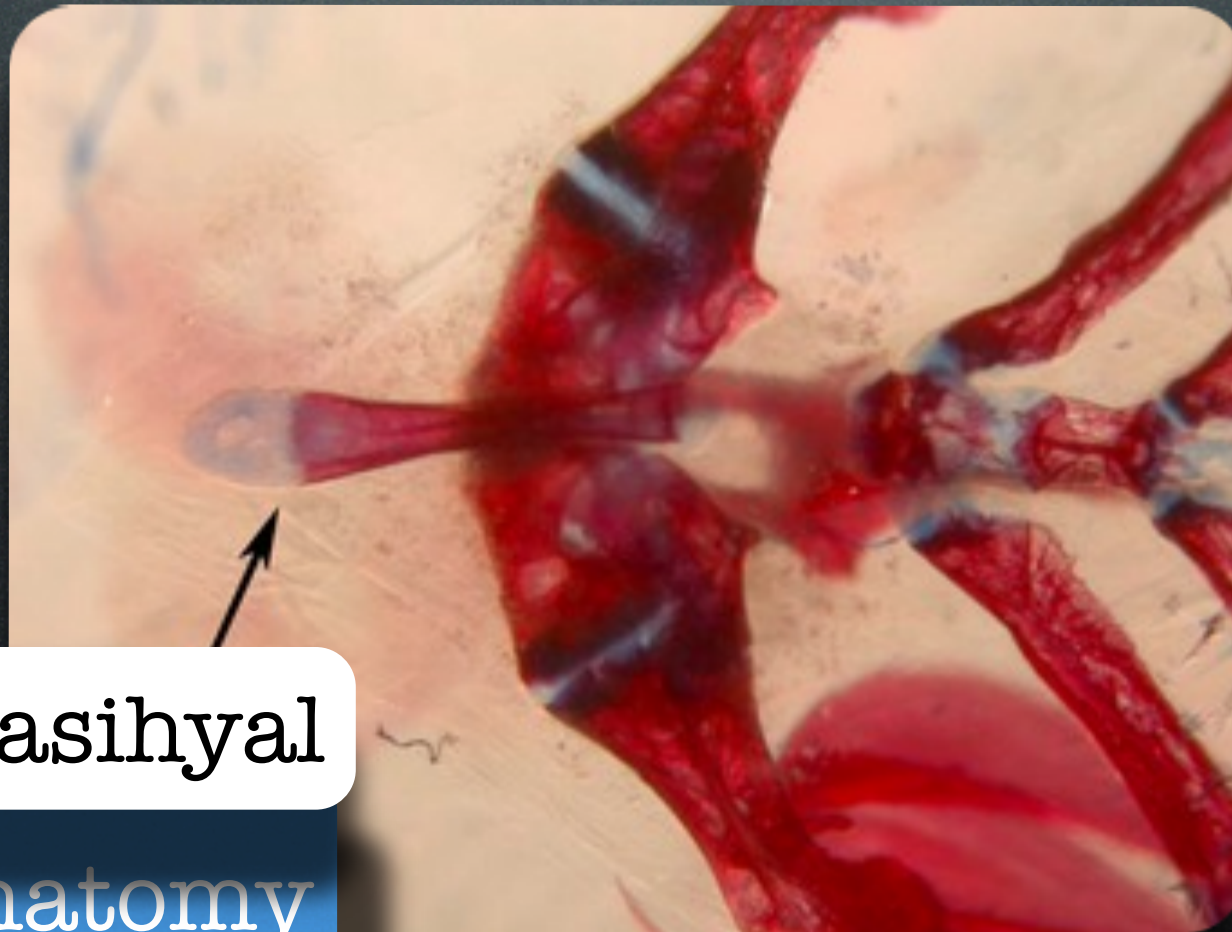
Requirements:

- Develop ontologies
- Annotate evolutionary phenotypes
 - using ontologies
 - using a syntax shared with model organism databases
- Develop curation software

Annotation

- Tagging or marking up something (image, text) with terms from one or more ontologies

Image annotated with ontology term



Basihyal

Teleost Anatomy
Ontology
TAO:0000316

Phenotypes annotated with ontology terms and EQ syntax

Entity

Caudal fin
ZFA: 0001058

Quality

Decreased size
PATO: 0000587



no tail
mutant

Teleost Anatomy Ontology: TAO

Phenotype and Trait Ontology: PATO

Annotation of evolutionary phenotypes

Free-text character:

“Character 1: Parietal and supraoccipital fused (state 1) or separated (state 2)”

Annotations:

Parietal fused_with supraoccipital

Parietal bone
TAO:0000486

Fused with
PATO:0000642

Supraoccipital bone
TAO:0000595

Parietal separated_from supraoccipital

Parietal bone
TAO:0000486

Separated from
PATO:0001505

Supraoccipital bone
TAO:0000595

Requirements:

- Develop ontologies
- Annotate evolutionary phenotypes
 - using ontologies
 - using a syntax shared with model organism databases
- **Develop curation software**

Phenex: software for curation of evolutionary phenotypes

The screenshot displays the Phenex software interface with the following components:

- Characters**: A list of 19 character descriptions, with 'Epiphyseal bar' (character 17) selected.
- States for Character...**: A table showing the states for the selected character.
- Term Info**: Detailed information for the 'epiphyseal bar' term, including its ID, ontology, and definition.
- Phenotypes for State: 0 - present**: A table showing the phenotype for the selected state.
- Complete Ontology Tree View**: A hierarchical tree view of the ontology, with 'epiphyseal bar' highlighted.

Characters

Character Description
1 Degree of development of ventral diverging lamellae of
2 Foramen on posterior portion of lateral wing of mesethm
3 Presence and degree of development of posterior portio
4 Orientation of distal portion of lateral wing of mesethmoi
5 Form of contact between vomer and mesethmoid
6 Form of anterior projection of vomer overlapping posteri
7 Presence or absence of a dorsolateral expansion of vo
8 Form of contact between vomer and palatine
9 Form of anterior process of lateral ethmoid
10 Form of nasal
11 Presence or absence of antorbital
12 Snout length
13 Presence or absence of rhinosphenoid
14 Well-developed bony expansion on ventral margin of r
15 Presence or absence of a frontal fontanel
16 Presence or absence of a parietal fontanel
17 Epiphyseal bar
18 Dorsal portion of sphenotic spine
19 Orientation of sphenotic spine

States for Character...

Symbol	State Description
0	present
1	absent

Term Info

epiphyseal bar

Basic Info

Term: **epiphyseal bar**

ID: TAO:0001502

Ontology: teleost_anatomy

Definition: Bilaterally paired medial bar of cartilage formed at midorbit through medial chondrification from triangles of cartilage that later form the junction of the taeniae marginales anterior, taeniae marginales posterior and the epiphyseal bar. Definition ref 1: ZFIN:curator

Links (1)

Parents

is_a [chondrocranium cartilage](#)

Children

DBxrefs (2)

Phenotypes for State: 0 - present

Entity	Quality	Related Entity	Comment
epiphyseal bar	present	None	None

Complete Ontology Tree View

- intestinal villus
- mesenchyme
- muscle
- nerve
- notochord
- portion of connective tissue
 - bone
 - cartilage
 - cranial cartilage
 - chondrocranium cartilage
 - auditory capsule
 - basal plate cartilage
 - epiphyseal bar
 - ethmoid cartilage
 - kinethmoid cartilage
 - lamina orbitonasalis
 - neurocranial trabecula
 - occipital arch cartilage
 - parachordal cartilage

Curation of phenotypic data

Taxon	#Pubs	#Characters	#Species	Annotations
Cypriniformes	10	725	930 (3,268)	674,250
Siluriformes	44	2582	2116 (2,867)	546,312
Characiformes	20	1661	595 (1,674)	988,295
Gymnotiformes	1	250	41 (134)	10,250
Gonorynchiformes	3	187	69 (37)	12,903
Clupeiformes	6	430	149 (364)	64,070
TOTAL	81	5960	4,000 (8,344)	7,225,780

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

Acknowledgements

- NSF DBI0641025 for Phenoscape funding (Mabee, Vision, Westerfield); Lapp, Dahdul, Kothari, Balhoff, Midford, Lundberg



- National Evolutionary Synthesis Center
NSF EF-0423641



- NIH HG002659 (to Monte Westerfield)



- Cypriniformes Tree of Life (NSF 0431290), colleagues and students (Mayden, Miya, Saitoh, He, Coburn, Arratia, Simon, Conway, Grey, Engeman, Bogutskya, Hilton, Aspinwall)



- Suzanna Lewis, Chris Mungall (Lawrence Berkeley National labs)

