PHENOTYPED

Paula Mabee

June 4, 2018

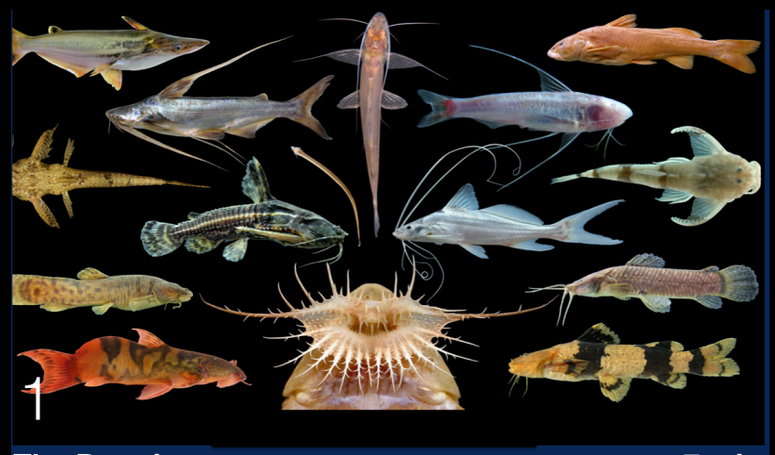
Second Annual Digital Data in Biodiversity Research Conference: Emerging Innovations for Biodiversity Data

University of California, Berkeley

UNIVERSITY OF SOUTH DAKOTA

NSF

PHENOTYPED

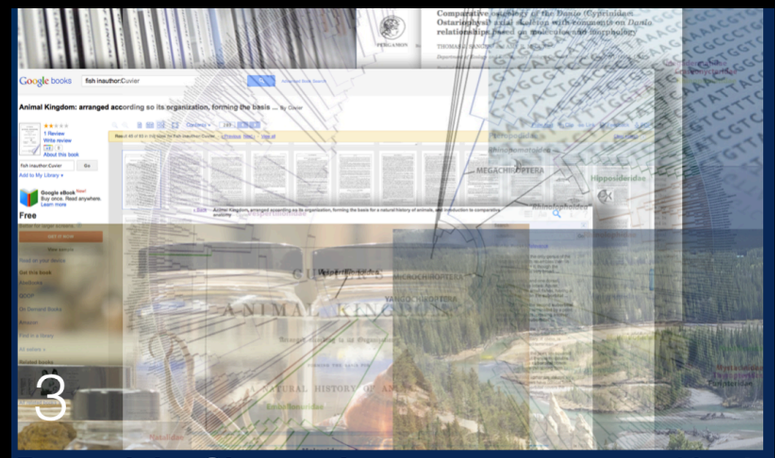


The Promise 7 min

Mabee and her colleagues at NESCent initiate a bootcamp with the MOD World. They fall for semantic phenotypes but are unaware of the perils ahead.

Missing Link 8 min

Moving beyond fundamentals, the scientists begin to use the System's logic on the Phenotyped. Dark secrets about the genetic bases of phenotypes come to light and others have ideas to reverse the logic.



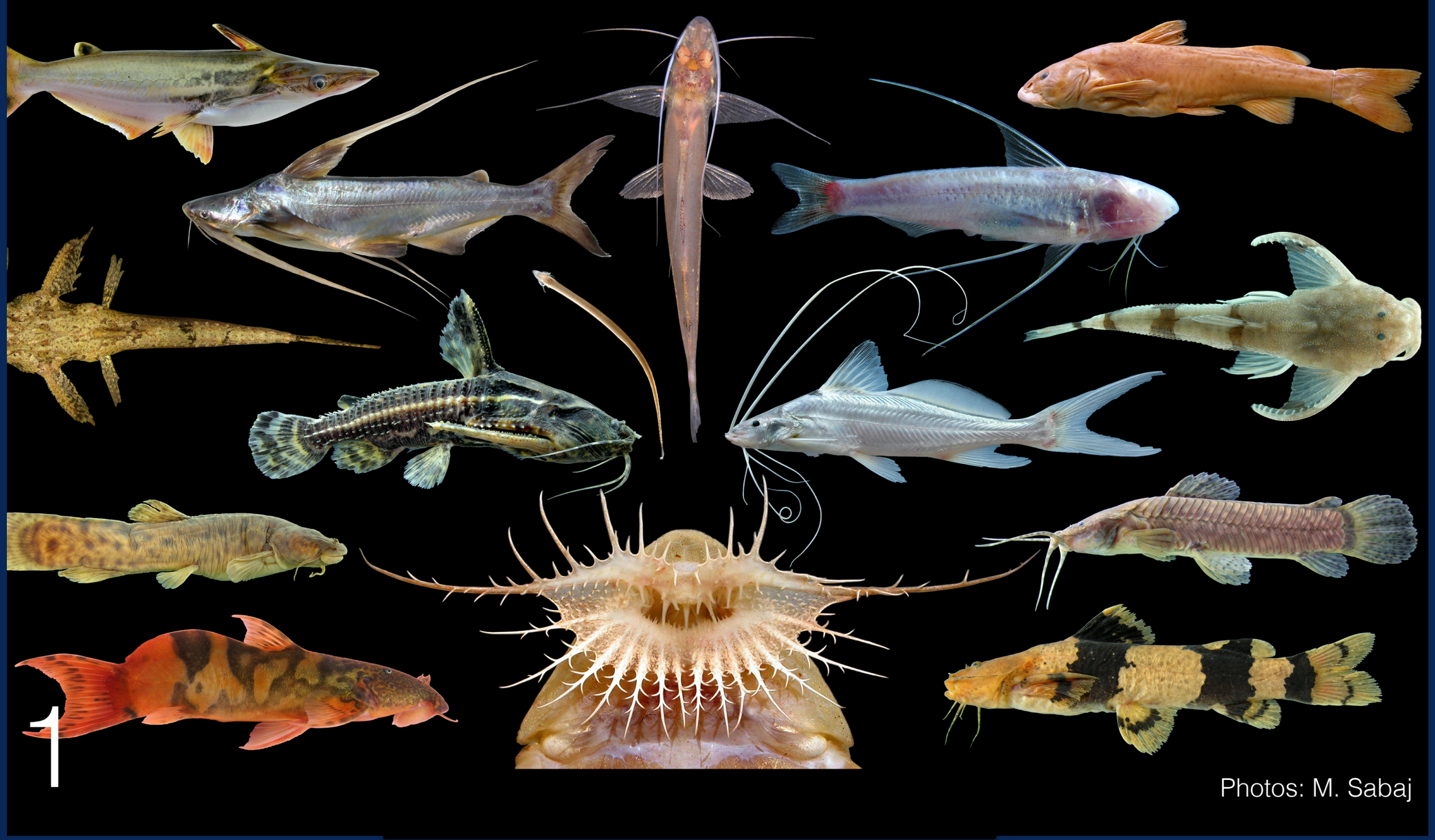
Beyond the Observed 10 min

When the scientists attempt to make a synthetic supermatrix of phenotypes, they realize that they can dramatically expand knowledge.

The Inter

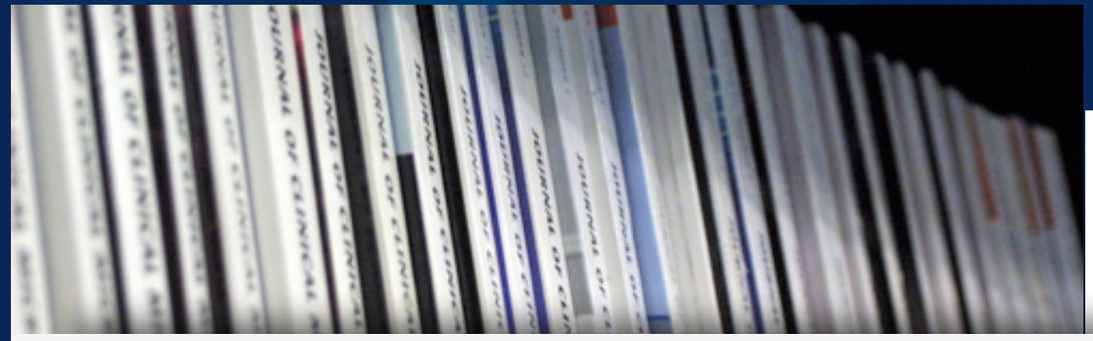
Worried h
in predict
phenotyp
which sp

OVERVIEW EPISODES TRAILERS & MORE MORE LIKE THIS DETAILS



The Promise 7 min

Mabee and her colleagues at NESCent initiate a bootcamp with the MOD World. They fall for semantic phenotypes but are unaware of the perils ahead.



PERGAMON

Zoological Journal of the Linnean Society, 2002, **135**, 529–546. With 14 figures

Comparative osteology of the *Danio* (Cyprinidae: Ostariophysi) axial skeleton with comments on *Danio* relationships based on molecules and morphology

THOMAS J. SANGER[1] and AMY R. McCUNE*

Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA

Google books fish inauthor:Cuvier Advanced Book Search

Animal Kingdom: arranged according so its organization, forming the basis ... By Cuvier

1 Review
Write review
+1 0
About this book

fish inauthor:Cuvier Go

Add to My Library

Google eBook New!
Buy once. Read anywhere.
Learn more

Free

Better for larger screens.

GET IT NOW

View sample

Read on your device

Get this book

AbeBooks
QOOP
On Demand Books
Amazon
Find in a library
All sellers »

Related books

All related books »

Contents 289

Plain text Clip Link Feedback PDF

Result **45** of **93** in this book for **fish inauthor:Cuvier** - ‹ Previous Next › - View all

Clear search

« Back **Animal Kingdom, arranged according so its organization, forming the basis for a natural history of animals, and inroduction to comparative anatomy** – Cuvier

CUVIER'S

ANIMAL KINGDOM,

Arranged according to its Organization;

FORMING THE BASIS FOR

A NATURAL HISTORY OF ANIMALS,

AND

Search

suborbital Go

Sort by: **Pages** | Relevance

The Star-gazer is the only genus of the Perch family which resembles them in this respect ; but in it, though the **suborbital** bone is very broad, ... 294

... small ventrals, and one dorsal, consisting of three lobes. Aputet, Treacherous, are small fishes, having a formidable spine on the **suborbital** ... 295

ScolojuUa, have the second **suborbital** plate toothed, and terminated by a point directed backwards, crossing another point of the third **suborbital**, ... 296

The second, with the characters of Capito, have the **suborbital** notched, showing the maxillary. if. chclo, is common in the Mediterranean ... 305

In the living state, the jaws are covered with fleshy lips, but there is no double lip adhering to the **suborbital** bones. These fishes have the oblong form ... 311

... the tongue and vomer are always smooth, but the jaws have conical teeth, and the large **suborbital** covers the cheek like an operculum. ... 319

ix / 676

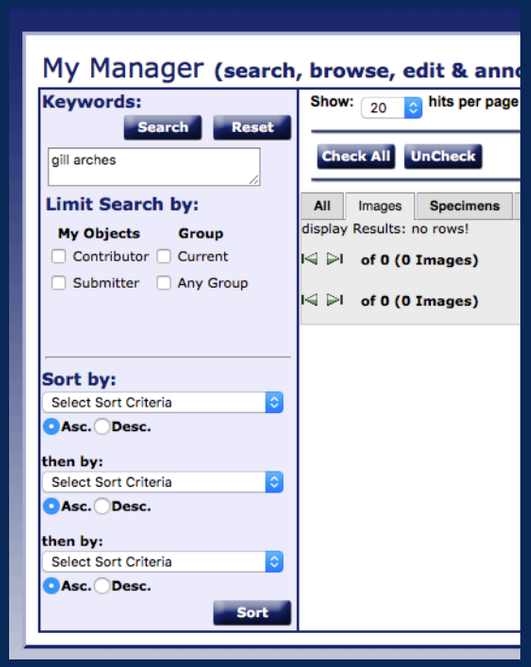
Rich legacy, digital, but not computable

Difficult for machines….

OMIM Query	# of records
“large bone”	785
"enlarged bone"	156
"big bones"	16
"huge bones"	4
"massive bones"	28
"hyperplastic bones"	12
"hyperplastic bone"	40
"bone hyperplasia"	134
"increased bone growth"	612

OMIM
Online Mendelian Inheritance in Man
Johns Hopkins University

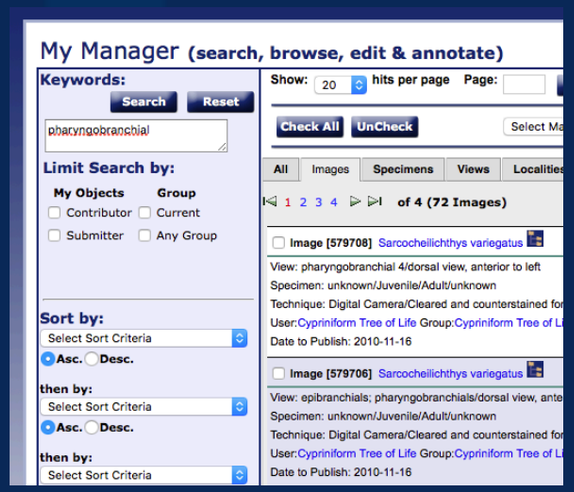
Discovery impeded



"*Find all images of keywords =* ***gill arch or gill arch skeleton***"

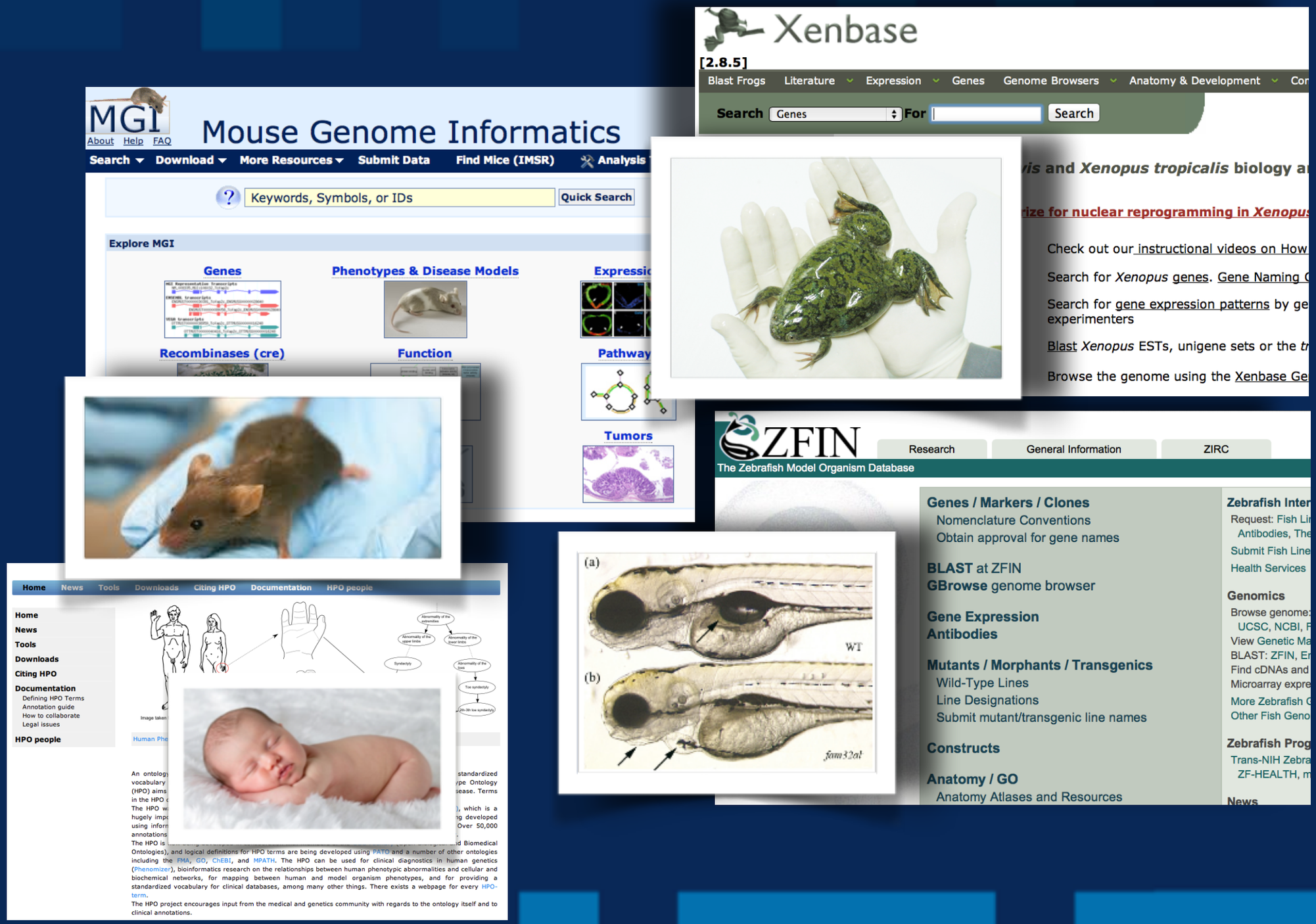
Results = 0

But there are many images of parts of gill arches

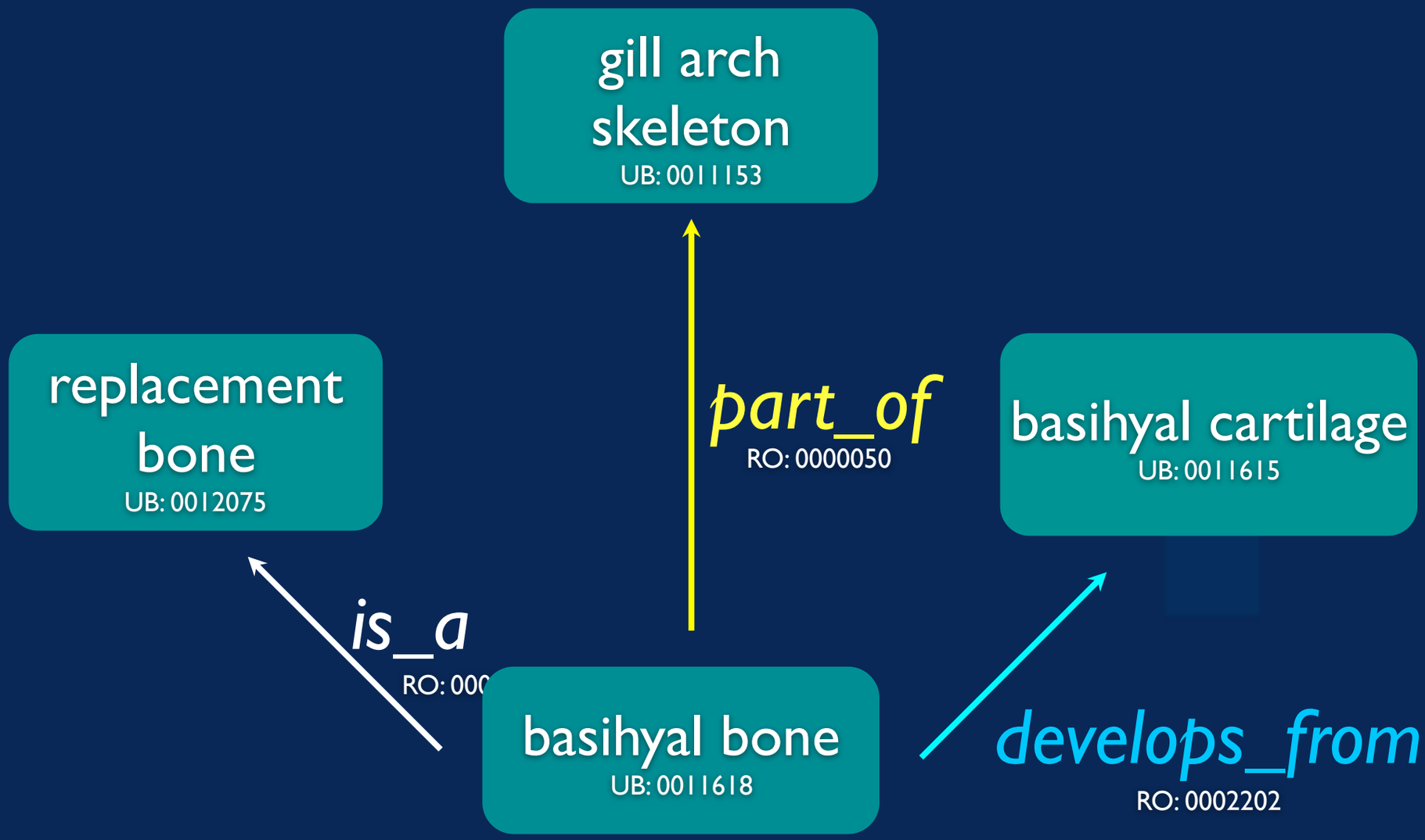


From Morphbank [28 May 2018]

MOD world: using ontologies, model organism databases link phenotypes—genes



Semantics provide meaning



part_of: http://purl.obolibrary.org/obo/BFO_0000050
develops_from: http://purl.obolibrary.org/obo/RO_0002202

Dahdul et al., 2010
Systematic Biology

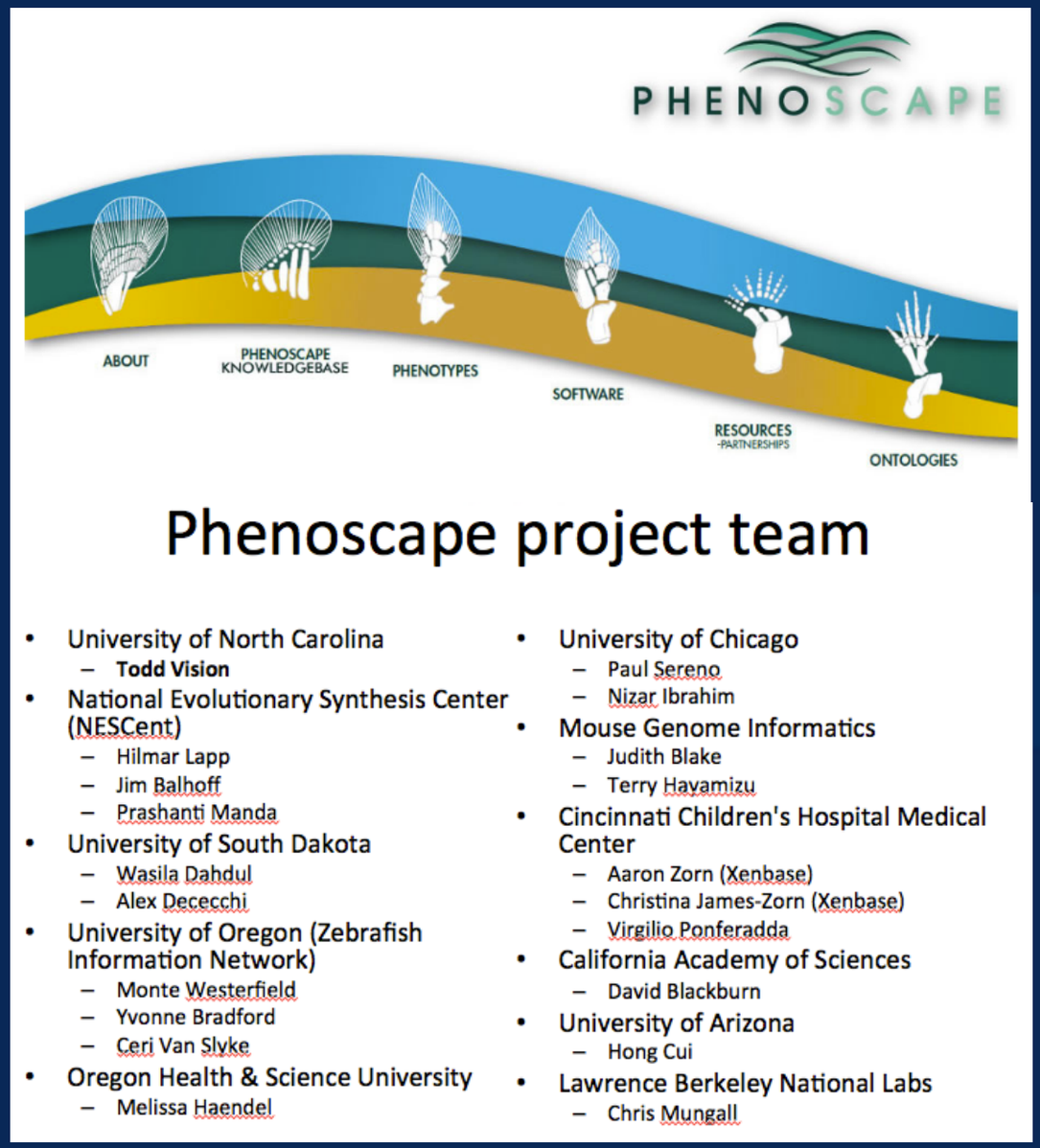
A (semantic) word is worth a thousand images…

Gill arch part	Number images
Basihyal bone	147
Basibranchial	236
Ceratobranchial	224
Hypobranchial	110
Epibranchial	220
Pharyngobranchial	92
Copula	70
TOTAL	**1099**

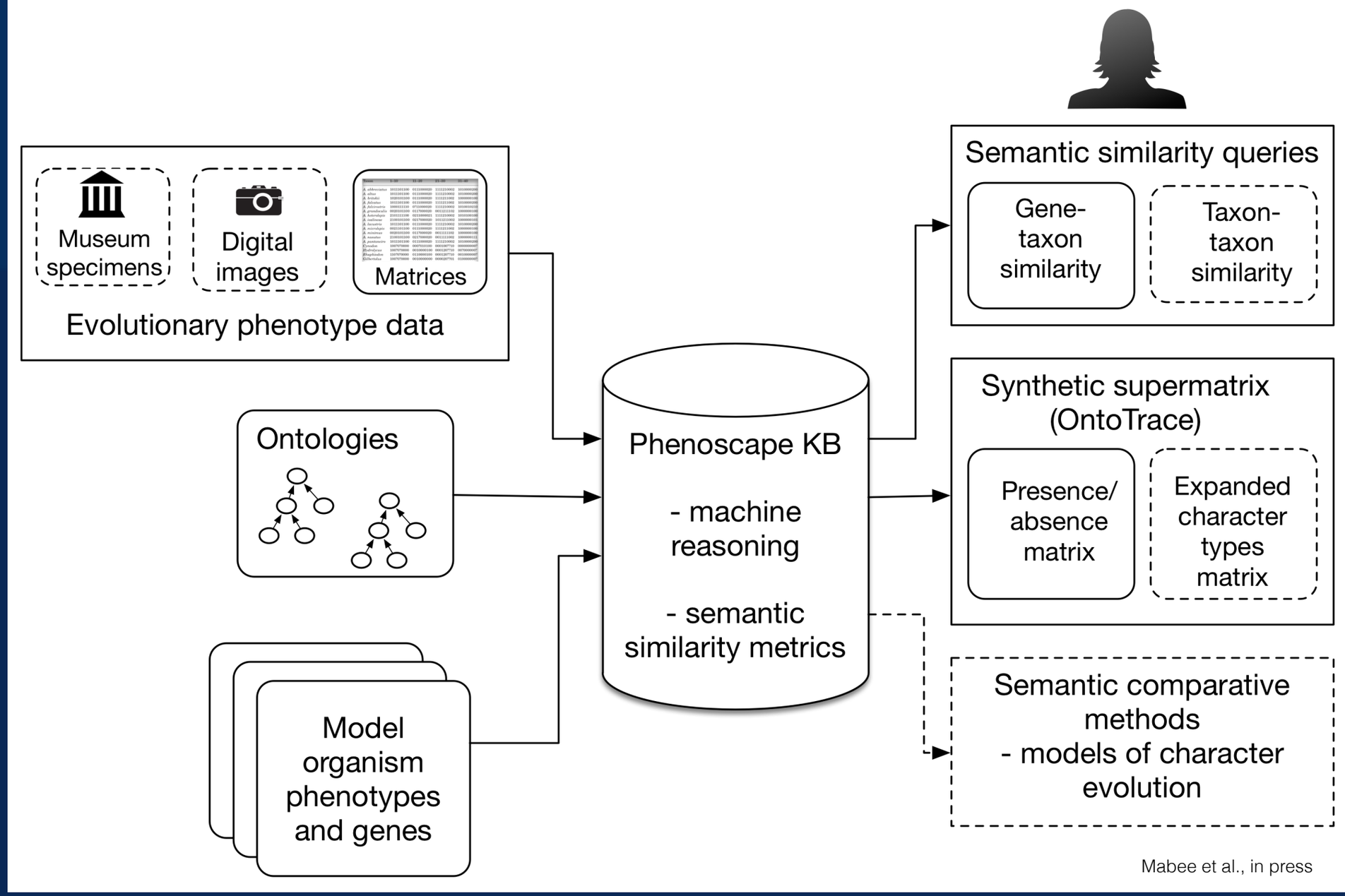
From Morphbank [28 May 2018]

Phenoscape (2007)

Promise: Find, compare, integrate computable phenotypes



phenoscape.org



Perils

Lots of resource & software development: phenotype ontologies, annotation software, Knowledgebase (KB), tools

Phenotyped: 4,260 species

but >1 billion biological specimens in U.S. natural history collections!

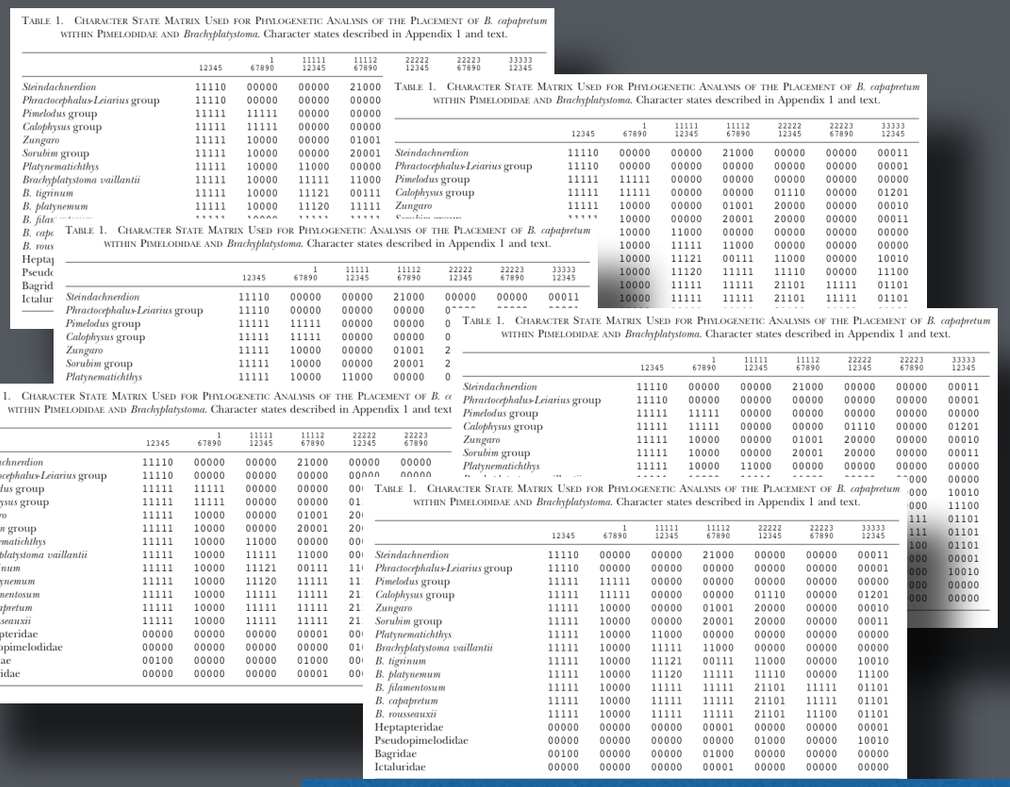
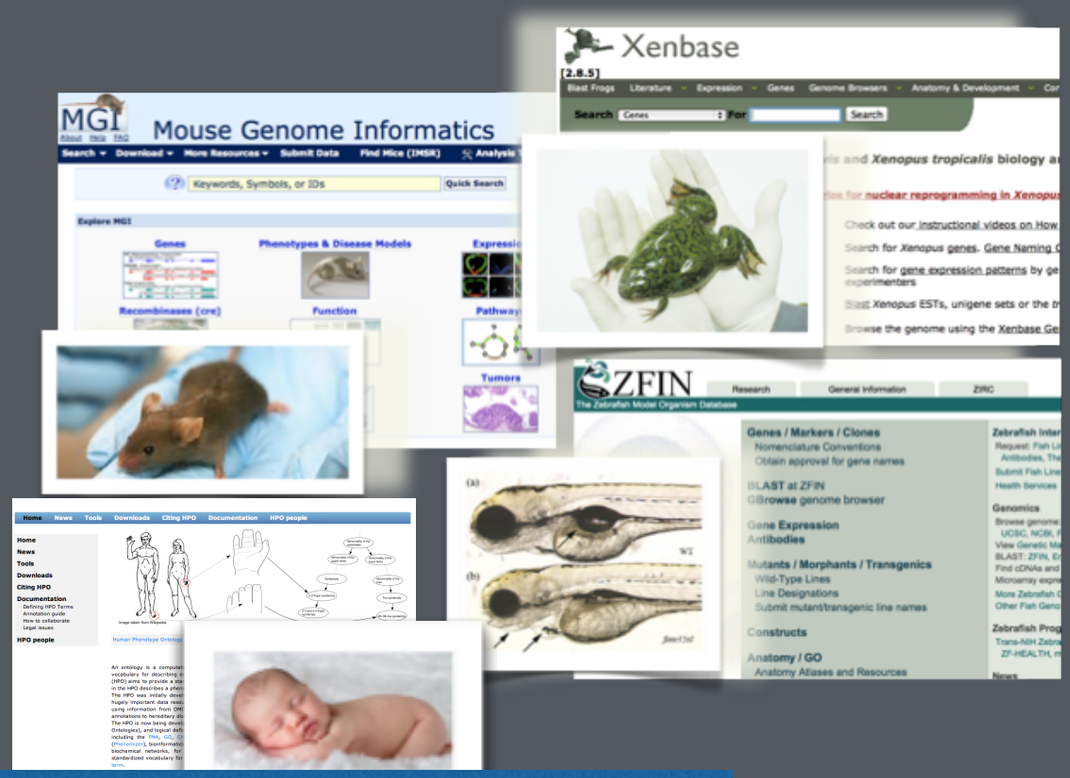


TABLE 1. CHARACTER STATE MATRIX USED FOR PHYLOGENETIC ANALYSIS OF THE PLACEMENT OF *B. capapretum* WITHIN PIMELODIDAE AND *Brachyplatystoma*. Character states described in Appendix 1 and text.

	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	01000
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

+



Phenoscape Knowledgebase

Anatomy studies

651,660 phenotypes

>5,000 taxa

161 studies

>20,000 character states

Model organism datasets

309,383 phenotypes

15,447 genes

+

732,658 anatomical expression

34,731 genes

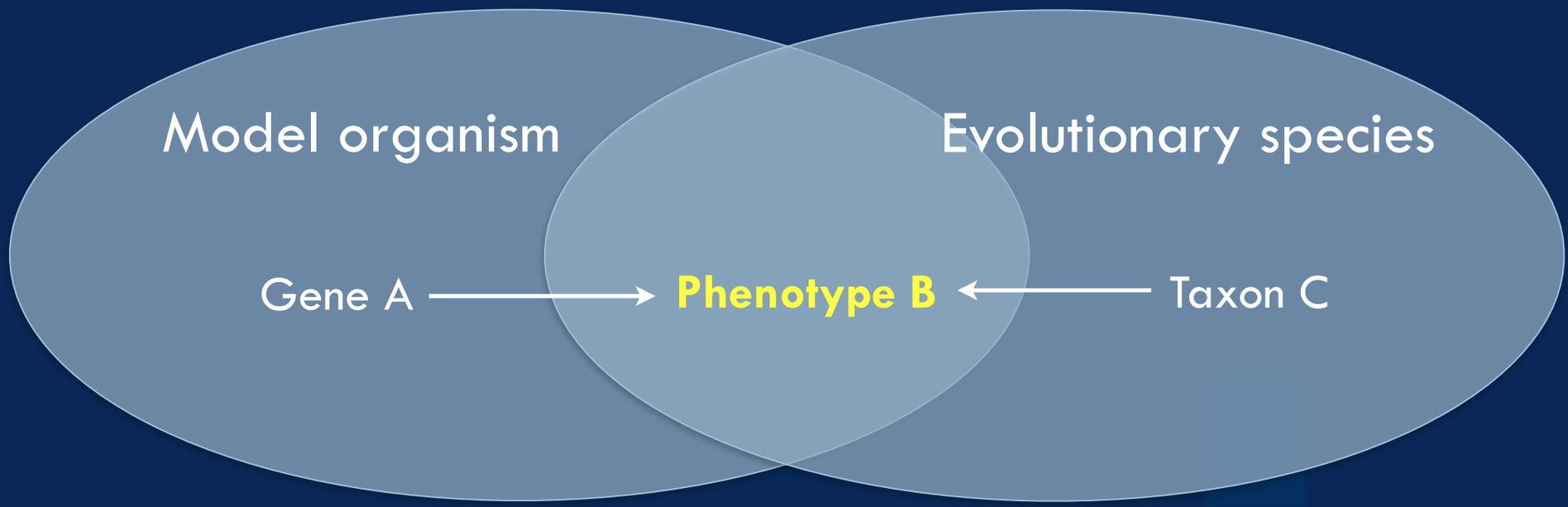
Phenotyped specimen data

- from legacy literature
- from images
- ‘born semantic’ phenotypes
- predicted phenotypes

Missing Link **8 min**

Moving beyond fundamentals, the scientists begin to use the System's logic on the Phenotyped. Dark secrets about the genetic bases of phenotypes come to light and others have ideas to reverse the logic.

Semantic phenotypes link species & genes



Gene-Phenotype connection

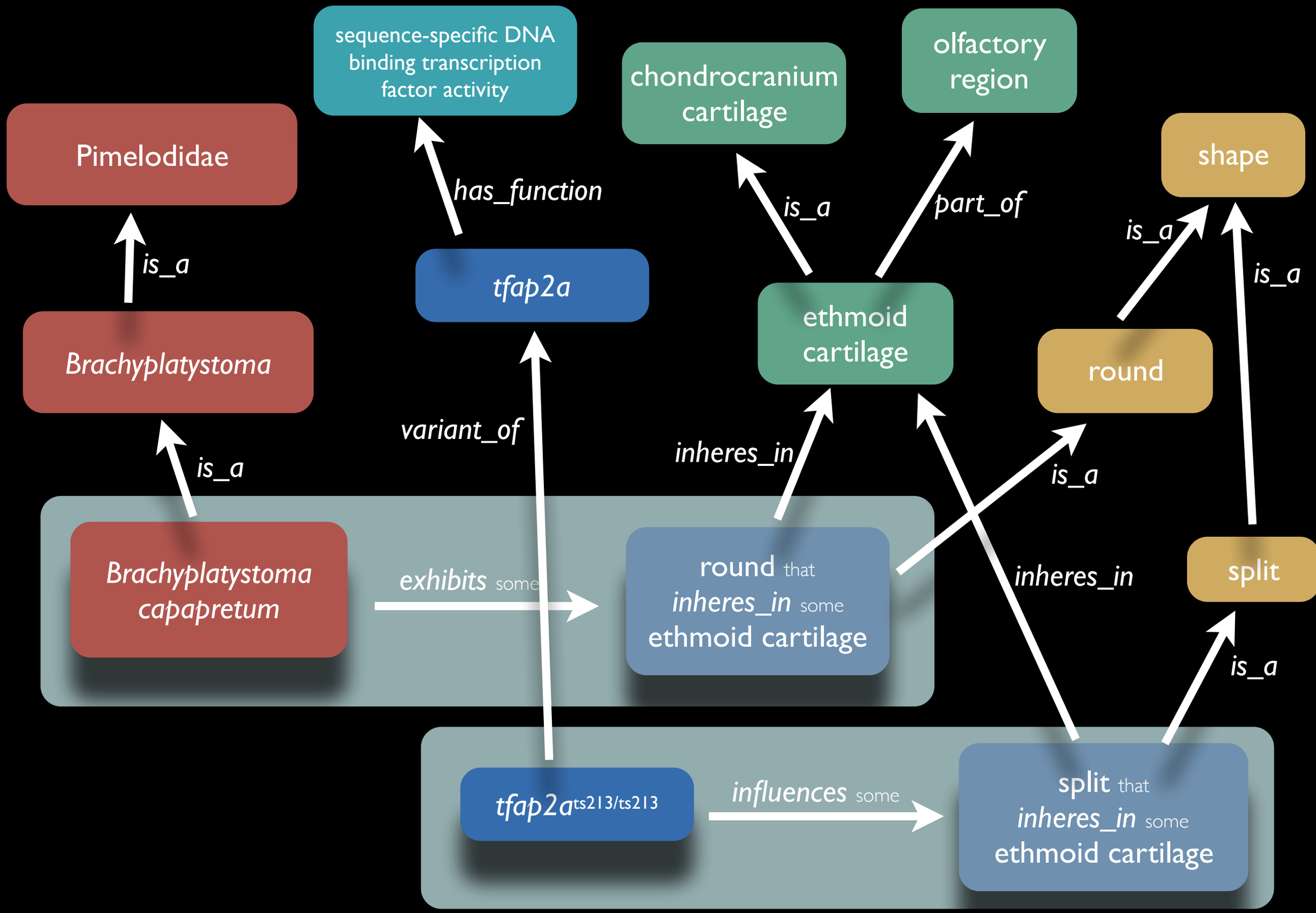


image:http://www.ustacrosport.com/buried-treasure

KB is a Hypothesis generator

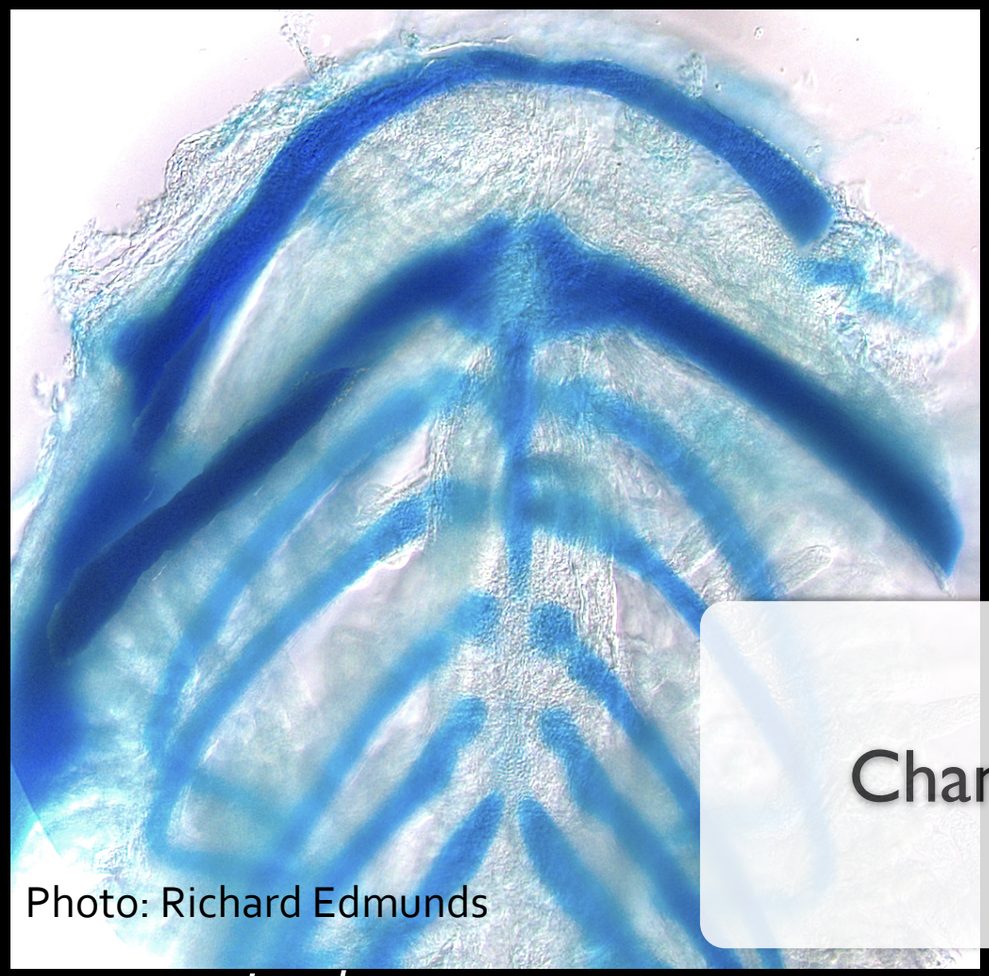
Interoperability yields 100,000's of Gene-Phenotype hypotheses

Mabee et al., 2012

E.g., How did the catfish lose its tongue?

Flathead Catfish; Photo by USFWS, used under Creative Commons License

Find the genes with a similar phenotype



Ictalurus punctatus

taxa: Siluriformes

Changes in *brpf1*?

Candidate genes

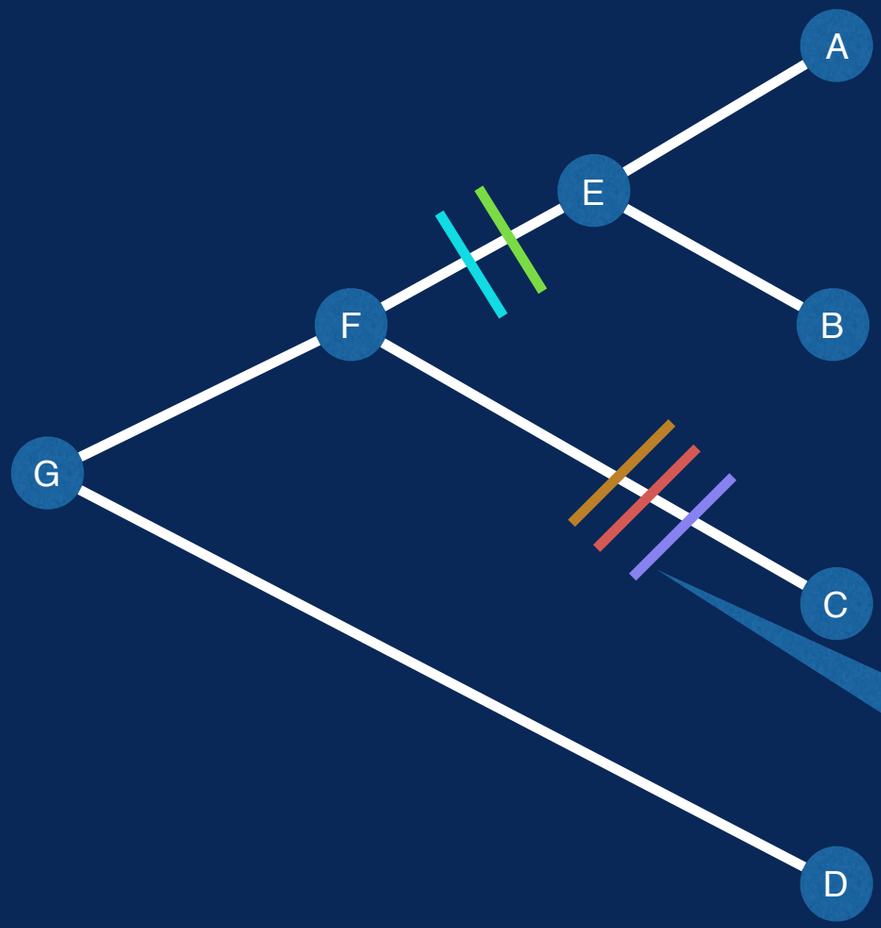
Laue *et al.*, 2008

genes: (11)

brpf1
disc1
disp1
fac
foxd3
hand2
myst3
sox9a
unm_th9
unm_tn20c
unm_ty5

Edmunds et al., 2016, Mol. Biol Evol.

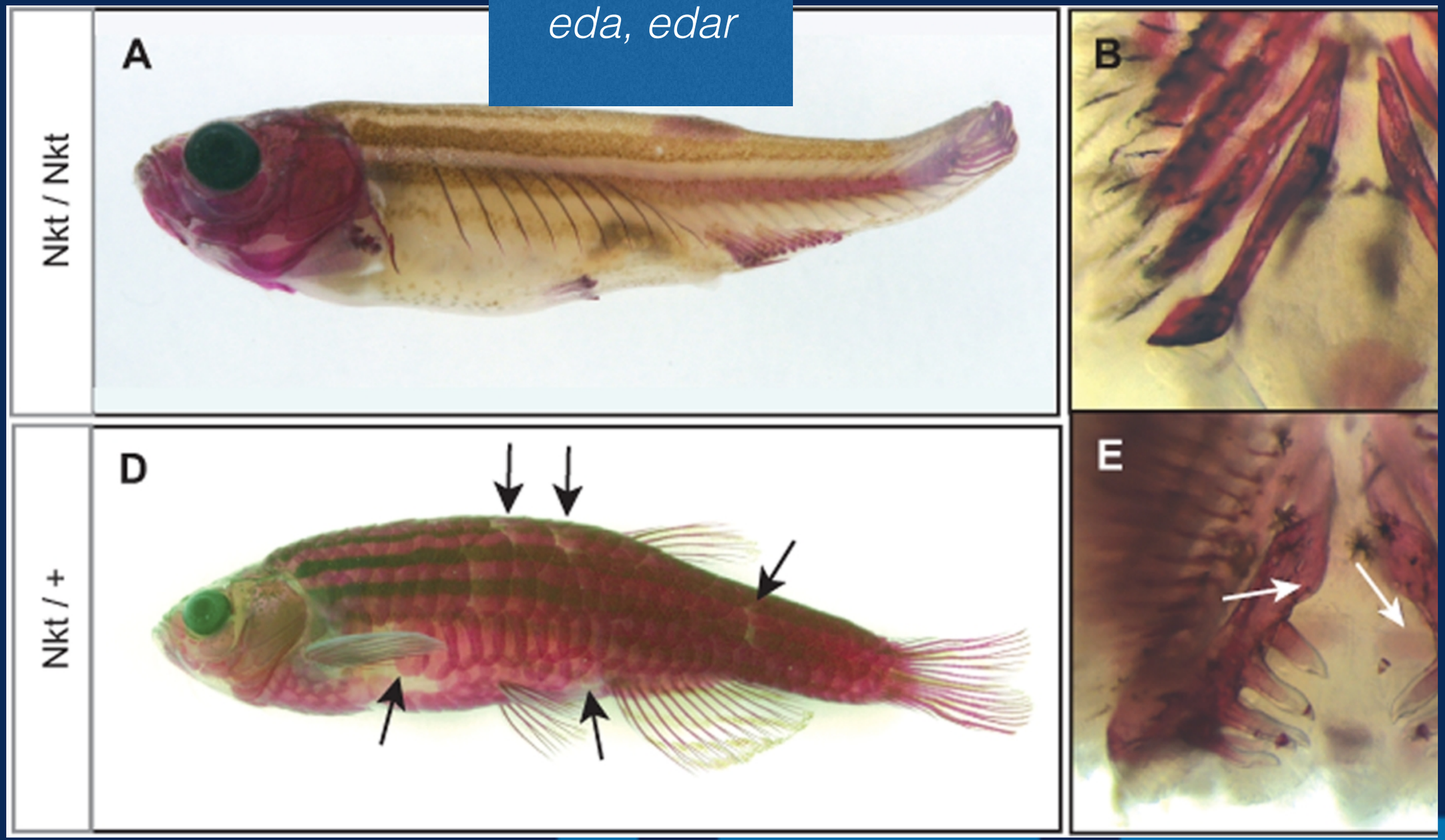
Phenoblast: find the genes with a similar set of phenotypes to a taxon



Set of phenotypes for a species

Reverse: Find the species with a phenotype similar to a gene…

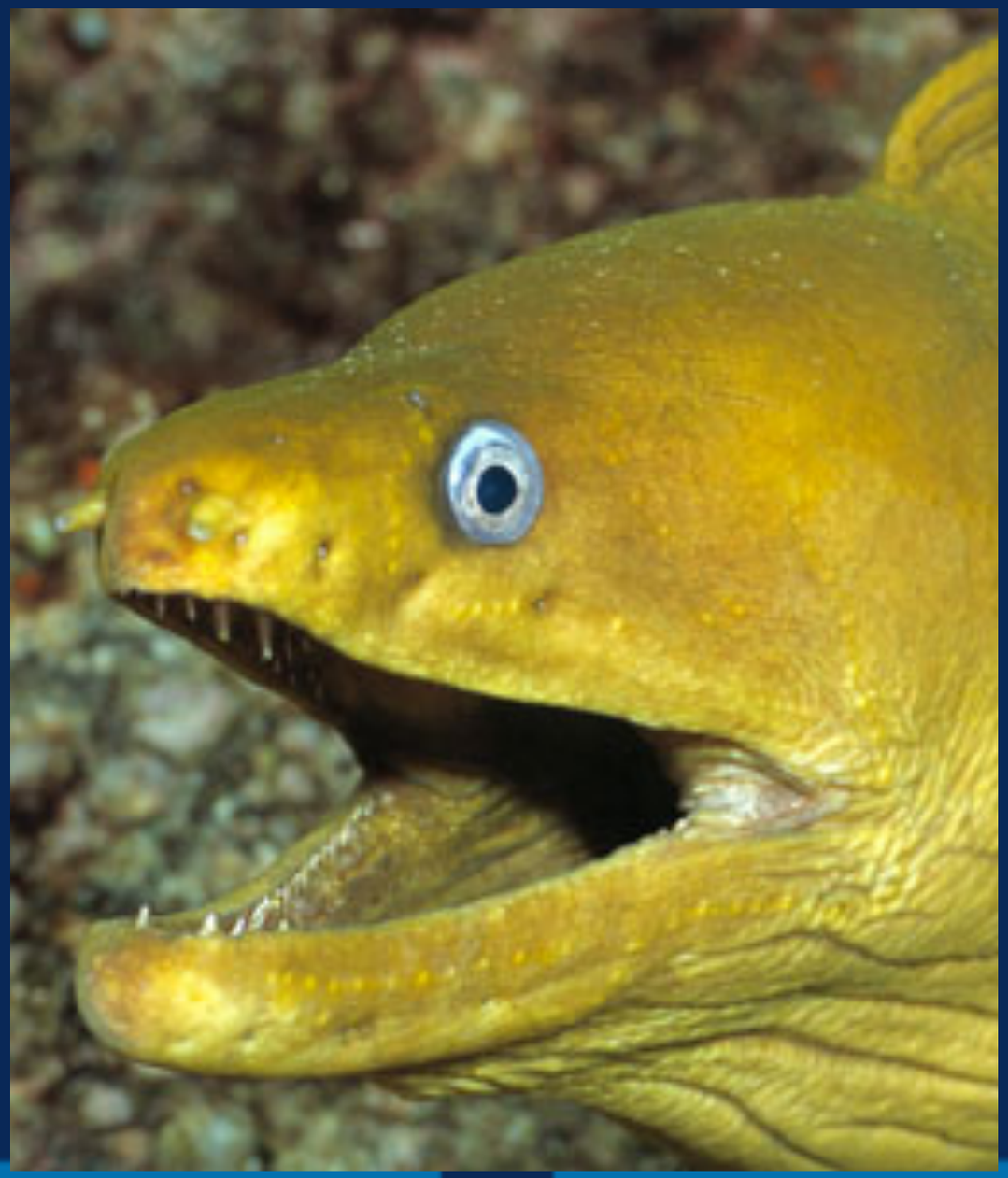
eda, edar



‘Candidate species’ (195)

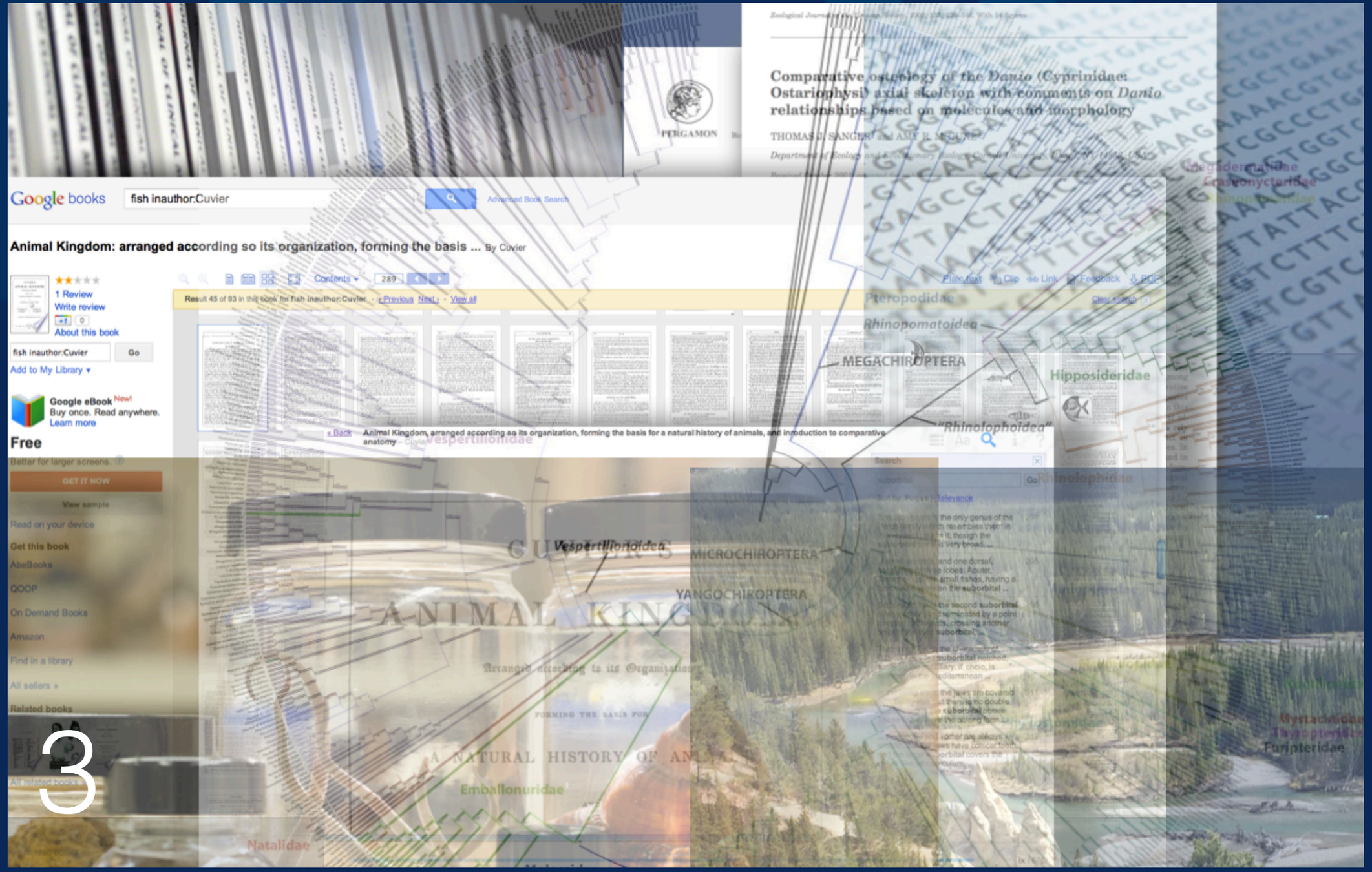
- including all eels, some catfish, etc.

- voucher specimens are in natural history collections



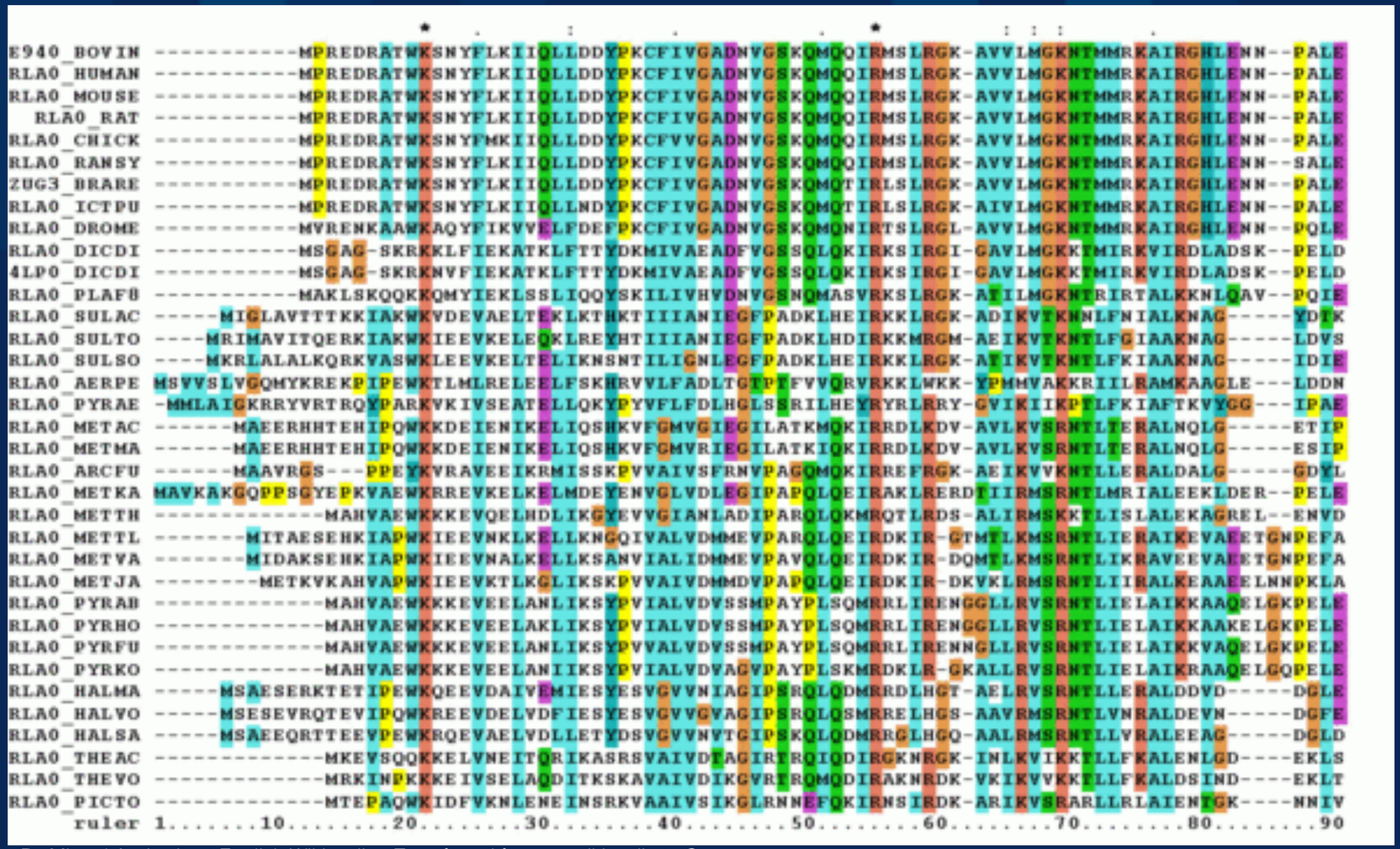
Phenotyped natural history specimens automatically linked with genetics

https://celebrating200years.noaa.gov/datasets/fishcollection/



Beyond the Observed 10 min

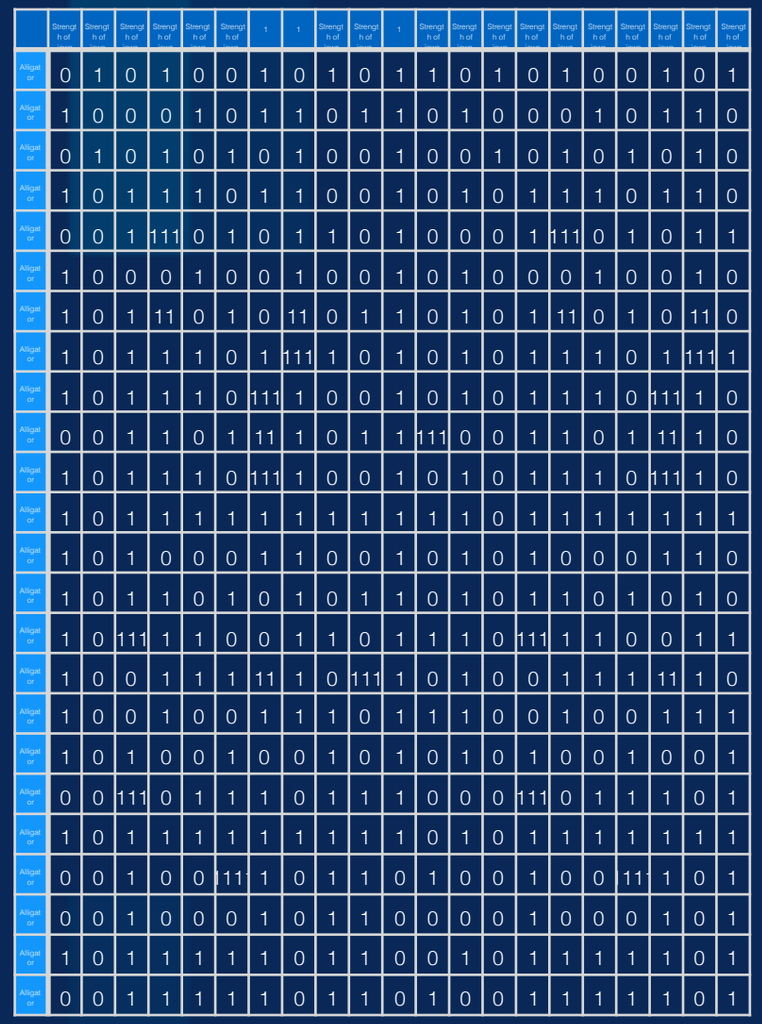
When the scientists attempt to make a synthetic supermatrix of phenotypes, they realize that they can dramatically expand knowledge.



Genetic supermatrix

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	10010
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



Phenotype supermatrix (+/-)

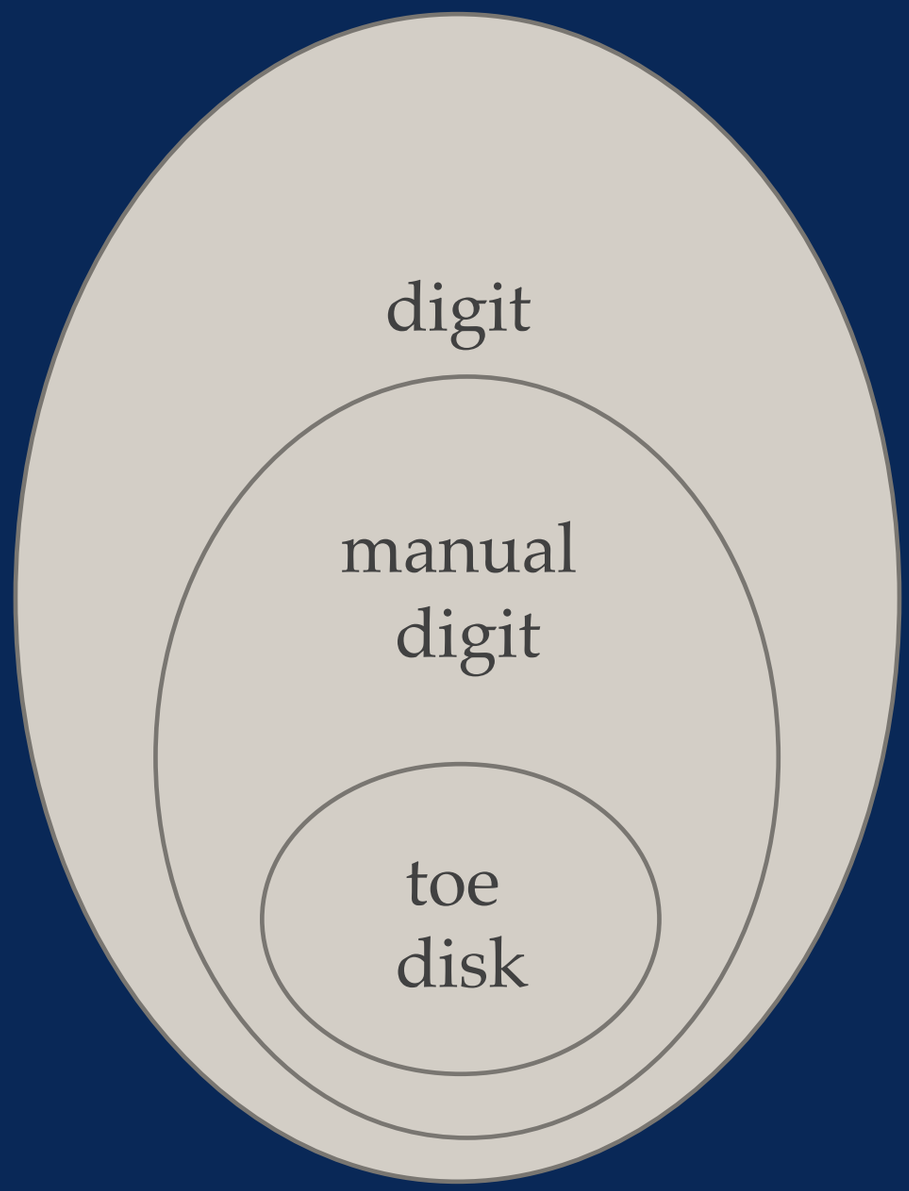
- ***OntoTrace***: Software tool that returns +/- synthetic matrices for desired taxa and characters

Dececchi, Balhoff, Lapp, & Mabee (2015)
Systematic Biology 64(6), 936-952.

Infer presence, from presence of part

- "Bottom up"
- If part of an entity is present, then the entity is present
- If a species has toe disks, it has digits

toe disks



digit is **present**

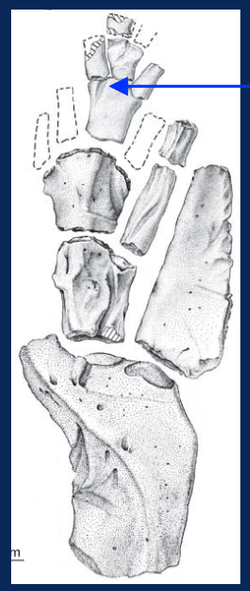
manual digit is **present**

toe disk is **present**
(Heyer, 1988)

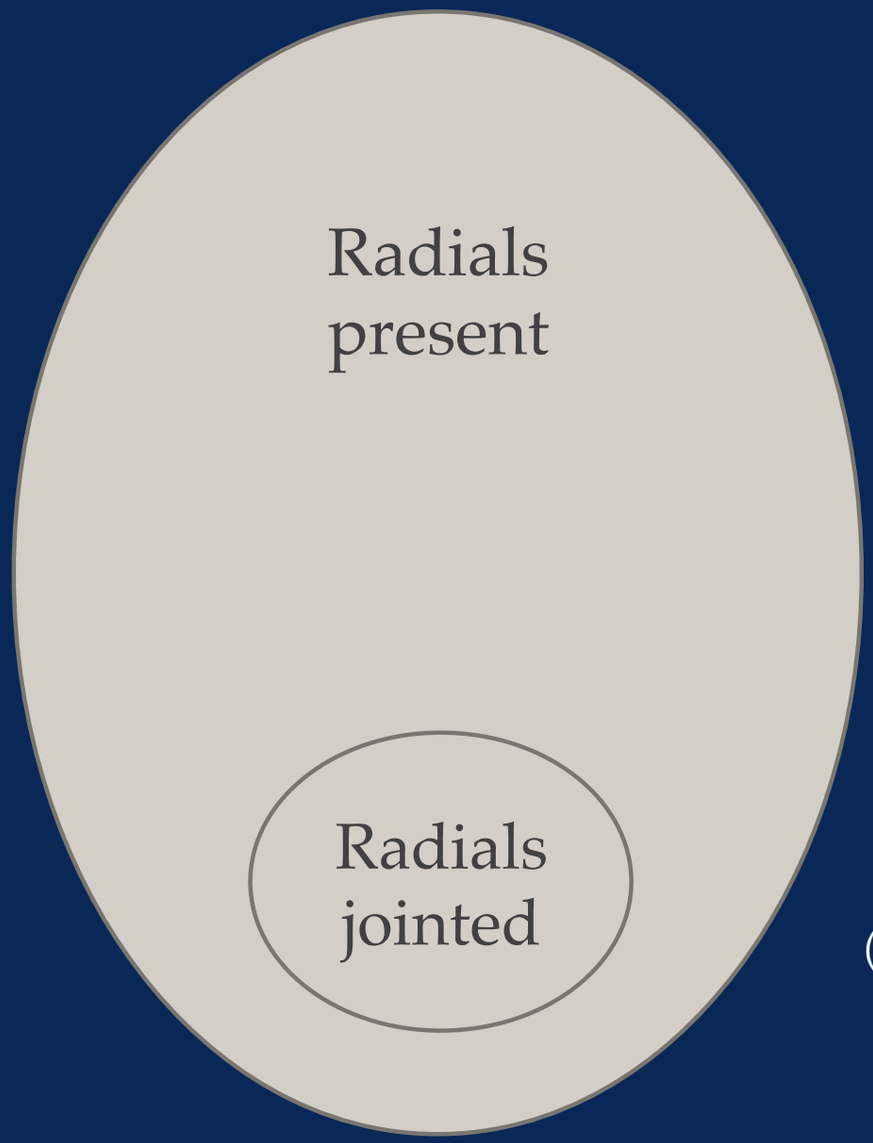
Infer presence, from quality of entity

- If an entity is asserted to have a quality other than absence, it is present

- E.g., If radials are jointed, they are present



Modified from Shubin et al. 2006

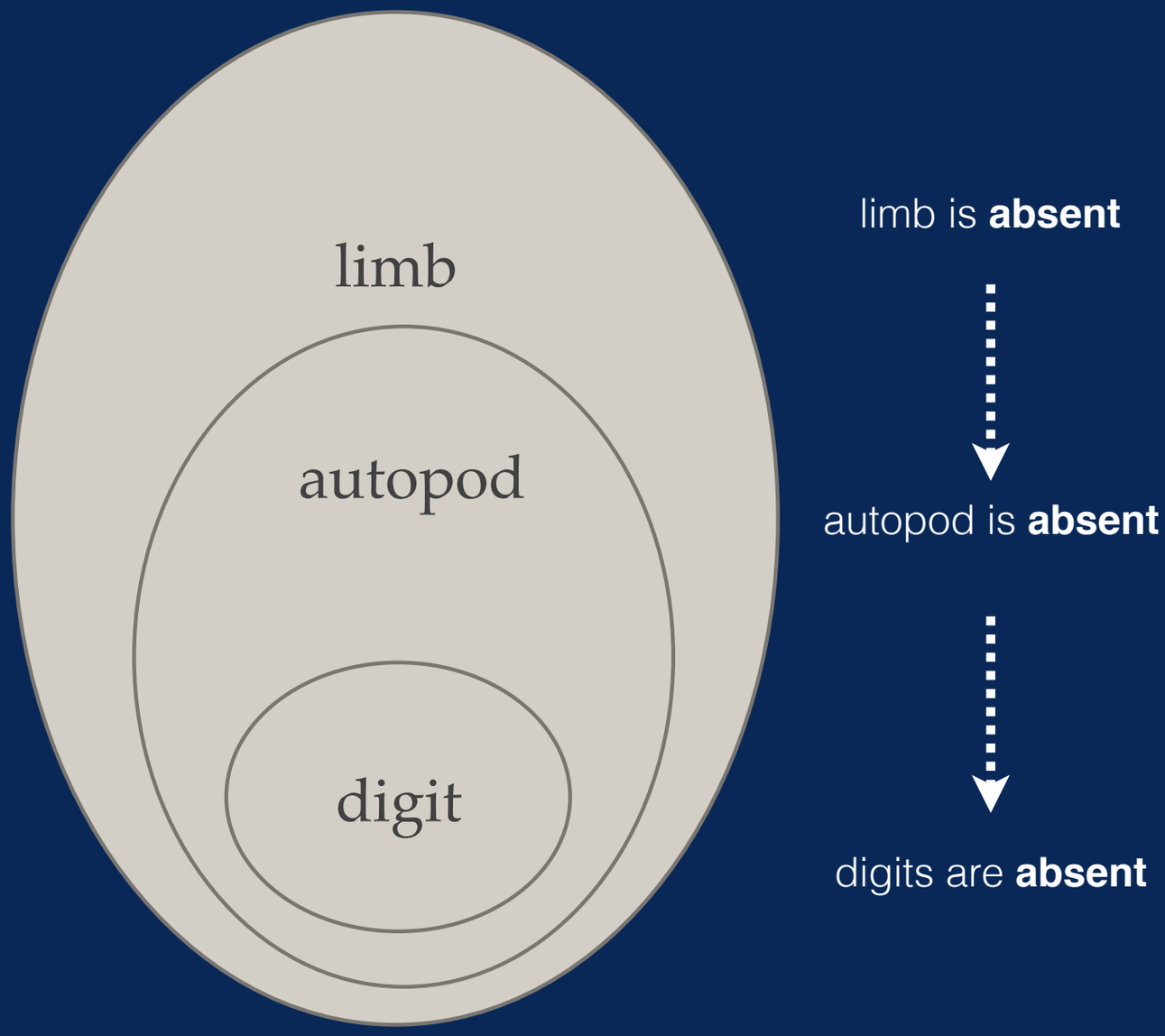


Radials are **present**

Radials are **jointed**
(Zhu & Ahlberg, 1998)

Infer absence, from absence of parent

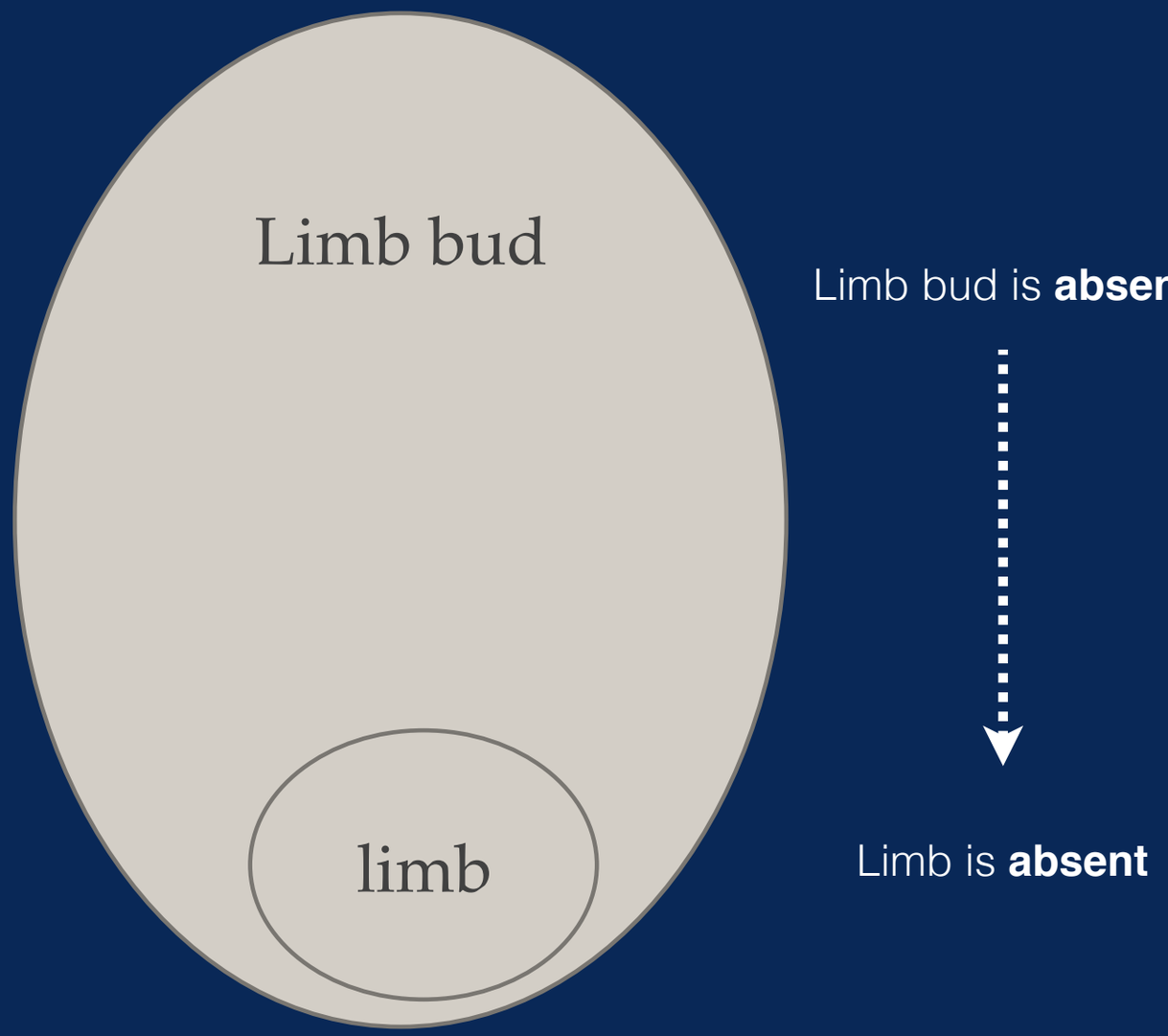
- “Top down”
- If parent is absent, all children and parts are absent
- If snakes lack limbs, they lack digits



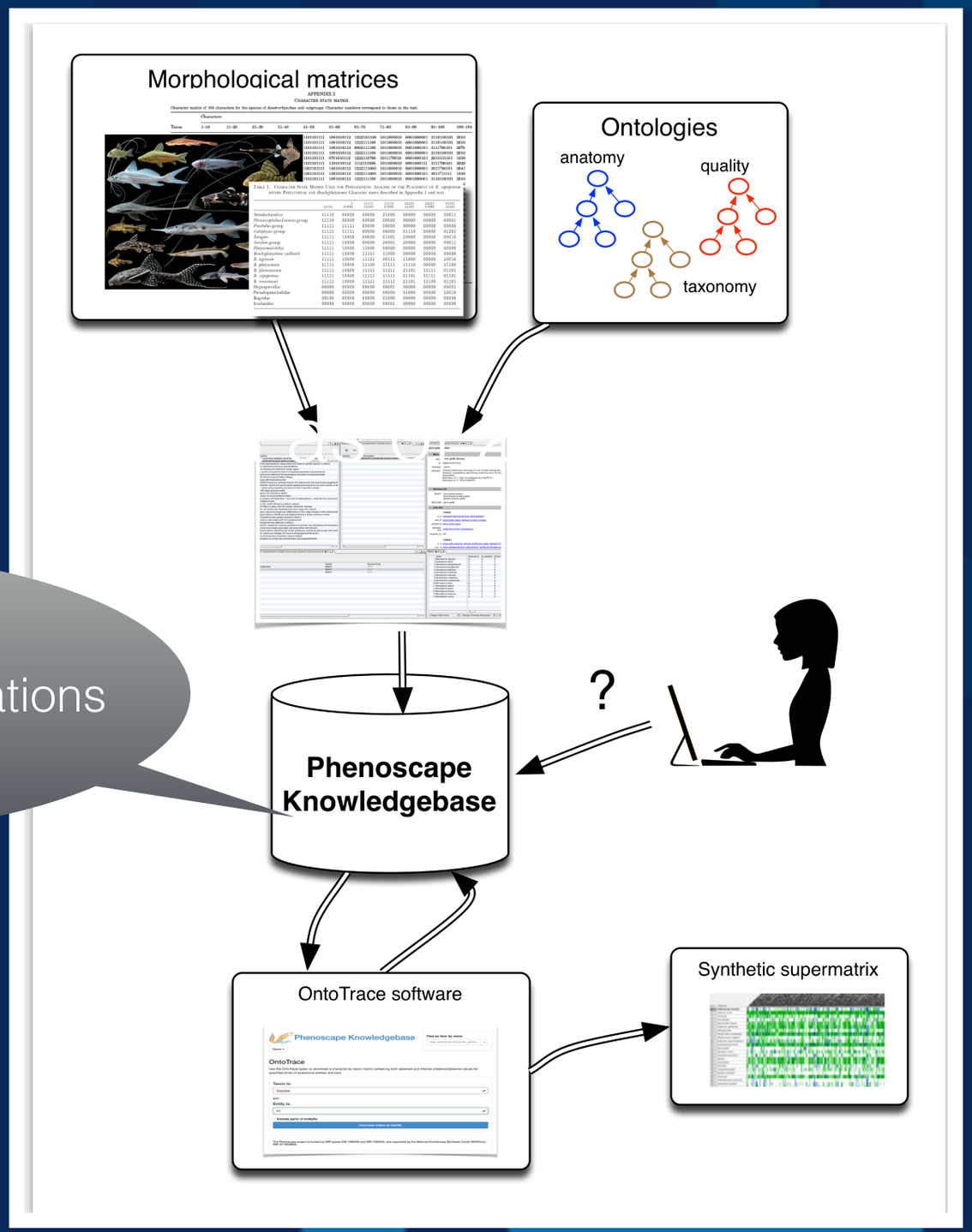
Infer absence, from absence of developmental precursor

- If developmental precursor is absent, entity is absent
- If species lacks limb buds, they lack limbs

from Sears, 2011



Query: “fin, limb, girdle & parts” “Sarcopterygii”

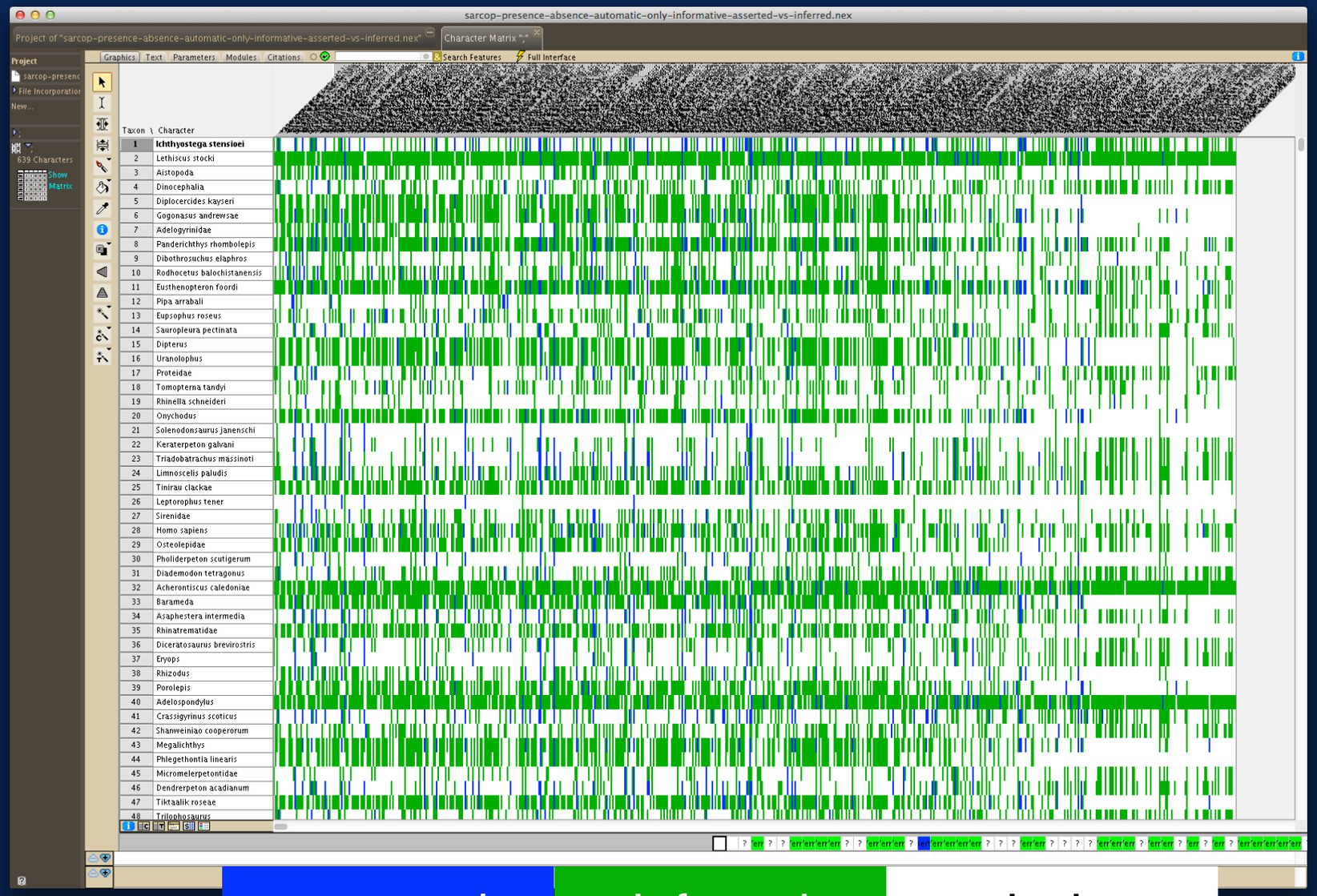


~150 studies
526K+ phenotype annotations
~5K vertebrates

(Data as of 2018-05-11)

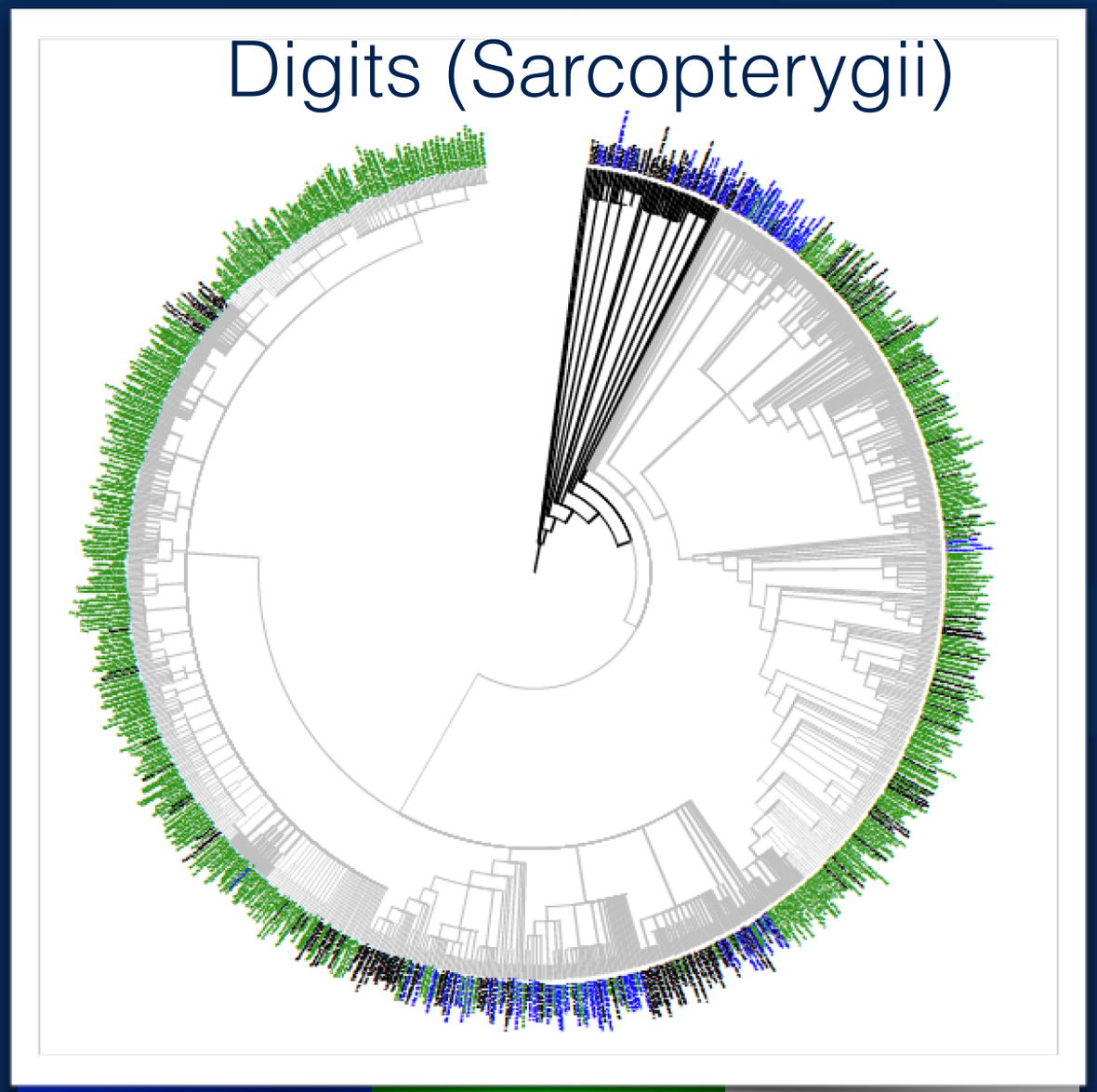
Dececchi, Balhoff, Lapp, & Mabee (2015) *Systematic Biology 64*(6), 936-952.

‘Synthetic morphological supermatrix’



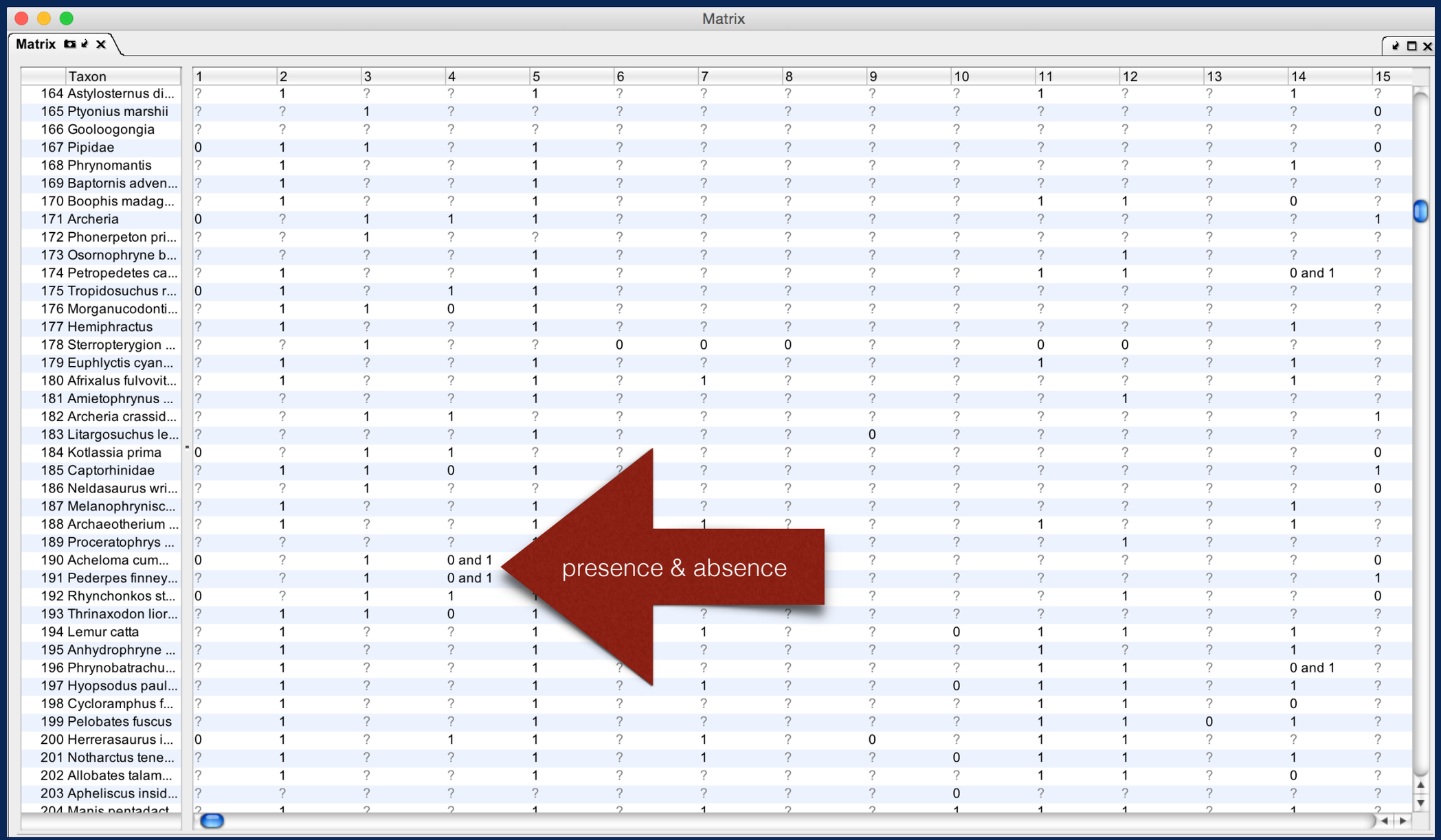
Dececchi, Balhoff, Lapp, & Mabee (2015) *Systematic Biology 64*(6), 936-952.

Inference provides most of the data for this phenotype



Dececchi, Balhoff, Lapp, & Mabee (2015) *Systematic Biology 64*(6), 936-952.

Enabled: Isolation, quantification of data conflicts

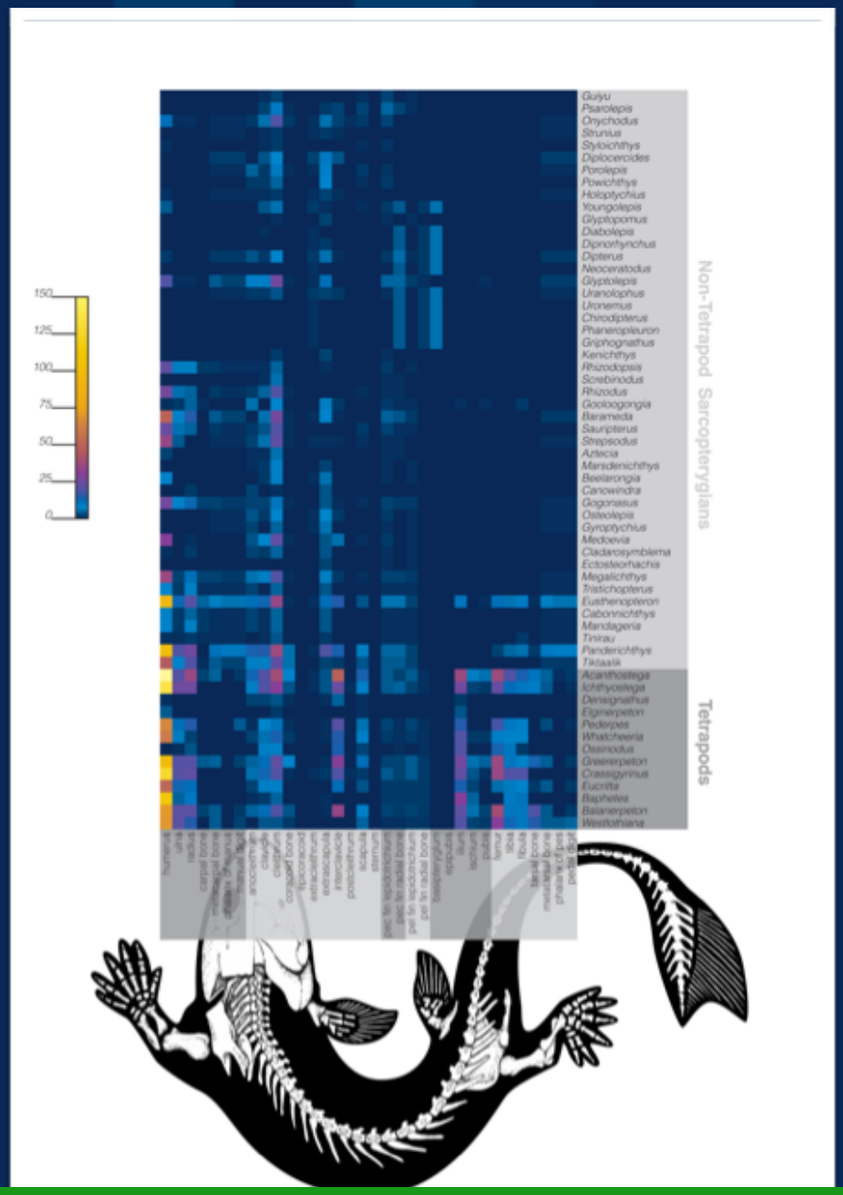


Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
164 Astylosternus di...	?	1	?	?	1	?	?	?	?	?	1	?	?	1	?
165 Ptyonius marshii	?	?	1	?	?	?	?	?	?	?	?	?	?	?	0
166 Gooloogongia	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
167 Pipidae	0	1	1	?	1	?	?	?	?	?	?	?	?	?	0
168 Phrynomantis	?	1	?	?	1	?	?	?	?	?	?	?	?	1	?
169 Baptornis adven...	?	1	?	?	1	?	?	?	?	?	?	?	?	?	?
170 Boophis madag...	?	1	?	?	1	?	?	?	?	?	1	1	?	0	?
171 Archeria	0	?	1	1	1	?	?	?	?	?	?	?	?	?	1
172 Phonerpeton pri...	?	?	1	?	?	?	?	?	?	?	?	?	?	?	?
173 Osornophryne b...	?	?	?	?	1	?	?	?	?	?	?	1	?	?	?
174 Petropedetes ca...	?	1	?	?	1	?	?	?	?	?	1	1	?	0 and 1	?
175 Tropidosuchus r...	0	1	?	1	1	?	?	?	?	?	?	?	?	?	?
176 Morganucodonti...	?	1	1	0	1	?	?	?	?	?	?	?	?	?	?
177 Hemiphractus	?	1	?	?	1	?	?	?	?	?	?	?	?	1	?
178 Sterropterygion ...	?	?	1	?	?	0	0	0	?	?	0	0	?	?	?
179 Euphlyctis cyan...	?	1	?	?	1	?	?	?	?	?	1	?	?	1	?
180 Afrixalus fulvovit...	?	1	?	?	1	?	1	?	?	?	?	?	?	1	?
181 Amietophrynus ...	?	?	?	?	1	?	?	?	?	?	?	1	?	?	?
182 Archeria crassid...	?	?	1	1	?	?	?	?	?	?	?	?	?	?	1
183 Litargosuchus le...	?	?	?	?	1	?	?	?	0	?	?	?	?	?	?
184 Kotlassia prima	0	?	1	1	?	?	?	?	?	?	?	?	?	?	0
185 Captorhinidae	?	1	1	0	1	?	?	?	?	?	?	?	?	?	1
186 Neldasaurus wri...	?	?	1	?	?	?	?	?	?	?	?	?	?	?	0
187 Melanophrynisc...	?	1	?	?	1	?	?	?	?	?	?	?	?	1	?
188 Archaeotherium ...	?	1	?	?	1	?	1	?	?	?	1	?	?	1	?
189 Proceratophrys ...	?	?	?	?	1	?	?	?	?	?	?	1	?	?	?
190 Acheloma cum...	0	?	1	0 and 1	1	?	?	?	?	?	?	?	?	?	0
191 Pederpes finney...	?	?	1	0 and 1	1	?	?	?	?	?	?	?	?	?	1
192 Rhynchonkos st...	0	?	1	1	1	?	?	?	?	?	?	1	?	?	0
193 Thrinaxodon lior...	?	1	1	0	1	?	?	?	?	?	?	?	?	?	?
194 Lemur catta	?	1	?	?	1	?	1	?	?	0	1	1	?	1	?
195 Anhydrophryne ...	?	1	?	?	1	?	?	?	?	?	1	?	?	1	?
196 Phrynobatrachu...	?	1	?	?	1	?	?	?	?	?	1	1	?	0 and 1	?
197 Hyopsodus paul...	?	1	?	?	1	?	1	?	?	0	1	1	?	1	?
198 Cycloramphus f...	?	1	?	?	1	?	?	?	?	?	1	1	?	0	?
199 Pelobates fuscus	?	1	?	?	1	?	?	?	?	?	1	1	0	1	?
200 Herrerasaurus i...	0	1	?	1	1	?	1	?	0	?	1	1	?	?	?
201 Notharctus tene...	?	1	?	?	1	?	1	?	?	0	1	1	?	1	?
202 Allobates talam...	?	1	?	?	1	?	?	?	?	?	1	1	?	0	?
203 Apheliscus insid...	?	?	?	?	?	?	?	?	?	0	?	?	?	?	?
204 Manis pentadact...	?	1	?	?	1	?	1	?	?	1	1	1	?	1	?

774 conflicted cells (0.5%) of 146,451 populated cells

Enabled: Phenotype enrichment

- Heat map:
  - Cell color reflects number of character states:
    - for each anatomical entity (column)
    - for each taxon (row)
- Dark blue = empty cells

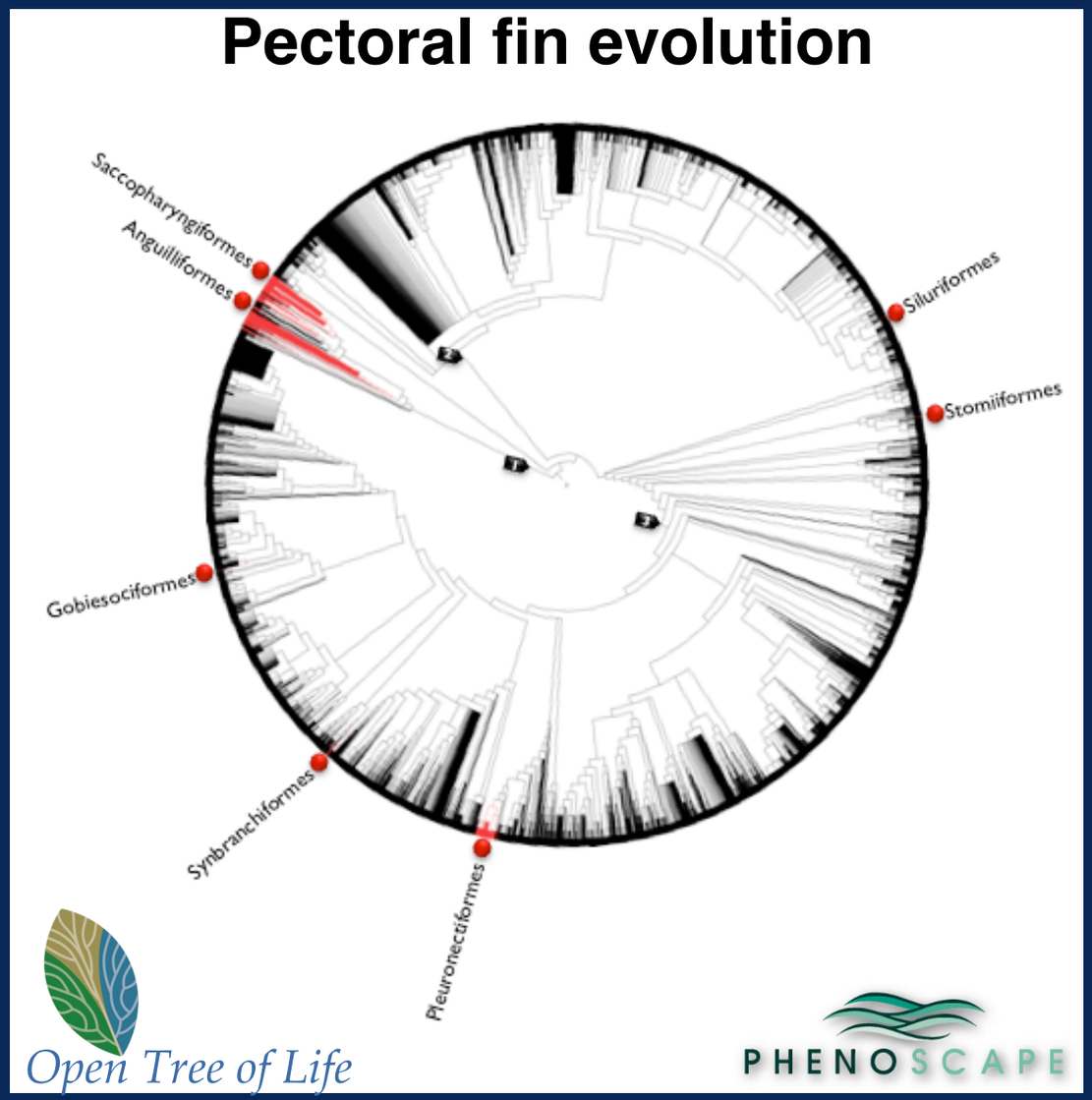


Sarcoptergyii, fins, limbs, girdles

Dececchi, Balhoff, Lapp, & Mabee (2015) *Systematic Biology 64*(6), 936-952.

Enabled: Large-scale ancestral state reconstruction

Enhancement to sparse data



Jackson, Balhoff, Hanscom, Franando, & Mabee (2018)
Systematic Biology doi.org/10.1093/sysbio/syx098

Potential: scale up phenotypic data to drive genomic discoveries

Cell Reports
Report

Cell PRESS
Open ACCESS

A "Forward Genomics" Approach Links Genotype to Phenotype using Independent Phenotypic Losses among Related Species

Michael Hiller,[1,5,*] Bruce T. Schaar,[1] Vahan B. Indjeian,[1,3] David M. Kingsley,[1,3] Lee R. Hagey,[4] and Gill Bejerano[1,2,*]
[1]Department of Developmental Biology
[2]Department of
Stanford Univ
[3]Howard Hugh
[4]Department of
[5]Present addr
01307 Dresde
*Corresponde
http://dx.doi.

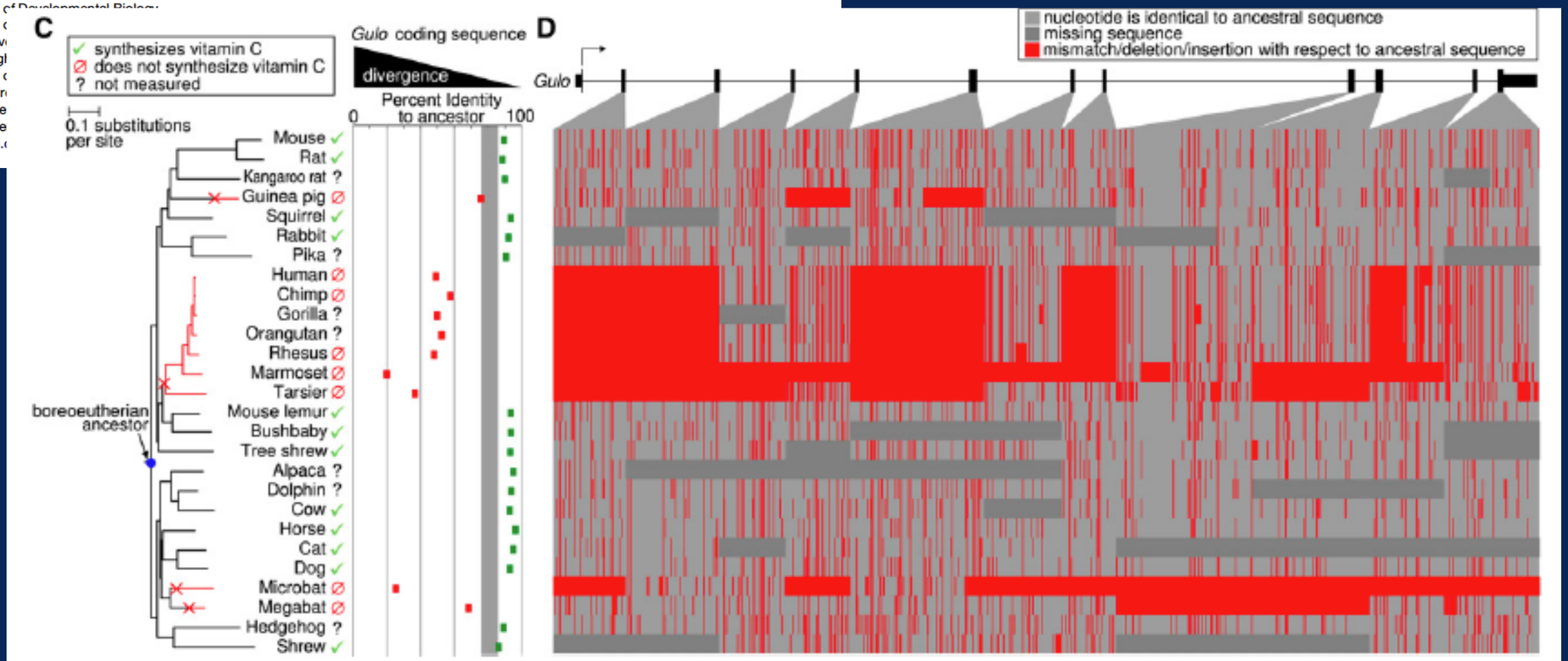


Figure 2. A Forward Genomics Screen to Match an Ancestral Presence/Absence Trait Pinpoints *Gulo* Inactivation in Vitamin C-Non-synthesizing Species

Hiller et al., 2012

Geography,
Environment

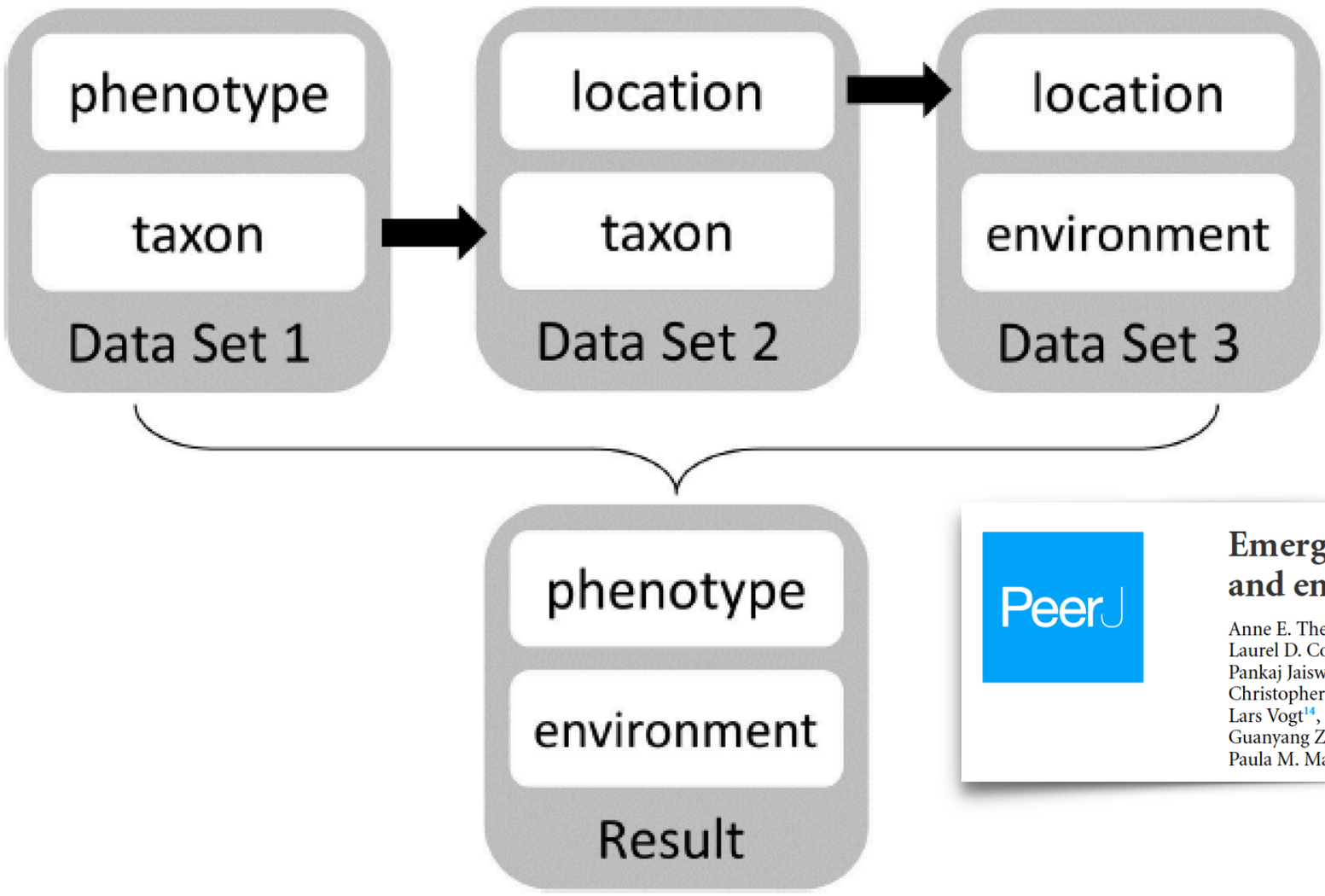
The Intermediary 3 min

Worried humans reached out to the Machine for help in predicting climate change. Cleverly connecting phenotype with collection data, It told the humans which species would be likely to win the competition.

PHENOTYPED

- Help me…
  - Related taxa with similar phenotypes from dissimilar environments
  - Unrelated taxa with similar phenotypes from similar environments
  - Understand the direction of phenotypic change in a particular species over time
  - Predict the likely changes in phenotype and species' interactions based on collections data

Reasoning across phenotype & environment

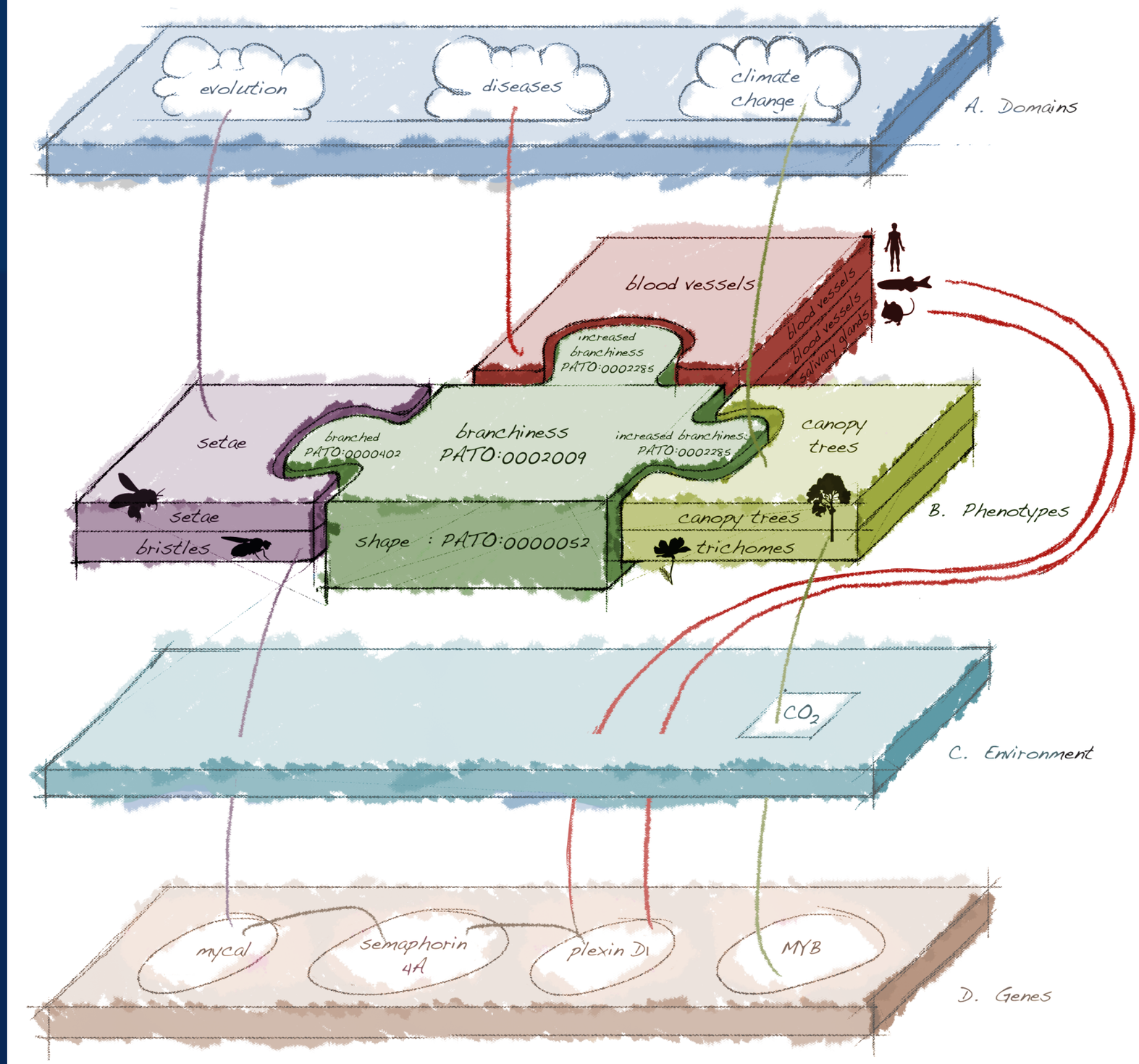


PeerJ

Emerging semantics to link phenotype and environment

Anne E. Thessen[1,2], Daniel E. Bunker[3], Pier Luigi Buttigieg[4], Laurel D. Cooper[5], Wasila M. Dahdul[6], Sami Domisch[7], Nico M. Franz[8], Pankaj Jaiswal[5], Carolyn J. Lawrence-Dill[9], Peter E. Midford[10], Christopher J. Mungall[11], Martín J. Ramírez[12], Chelsea D. Specht[13], Lars Vogt[14], Rutger Aldo Vos[15], Ramona L. Walls[16], Jeffrey W. White[17], Guanyang Zhang[8], Andrew R. Deans[18], Eva Huala[19], Suzanna E. Lewis[11] and Paula M. Mabee[6]

Figure 2 Manual workflow conceptual diagram. This diagram shows the manual workflow to link phenotype and environment data sets using current tools and services.



2015 Deans et al. *PLOS Biology*. January 5, 2015. doi: 10.1371/journal.pbio.1002033

Phenotypes a portal to understanding the integration between organism and environment

Acknowledgements

- The Phenoscape project is currently funded by NSF ABI Innovation collaborative grants (1661529, 1661356, 1661456, 1661516) and an ABI Development grant (1062542). Phenoscape was previously funded by NSF 1062404 and 0641025, and supported by NESCent, NSF 0905606.

- NSF Phenotype Research Coordination Network (phenotypercn.org), DBI-0956049

- Phenoscape contributors, Advisory Board, Data sources (see: http://phenoscape.org/wiki/Acknowledgments)

- Phenoscape collaborators