Propagation of Protein Properties: Principles and Practice

Mike Livstone

GO Annotation Camp

June 16, 2010

PAINT SOP

http://gocwiki.geneontology.org/index.php/RefG_Princeton_April_12-13_2010

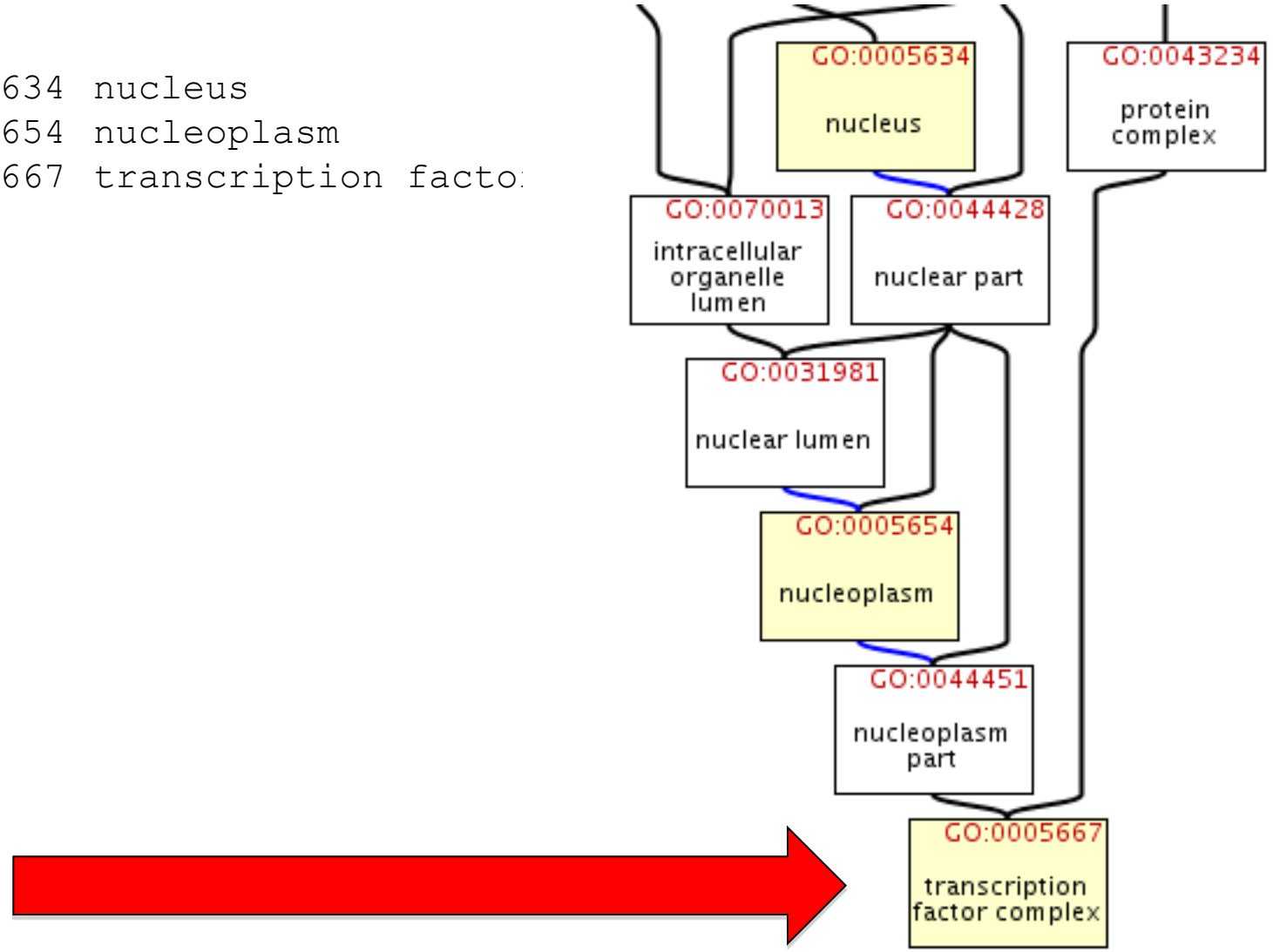
- Select the most specific GO term possible.
- Apply the term as widely as possible.
- When necessary, suggest changes to annotations or ontologies
- Use or address *every* experimental annotation.

Example: Forkhead/Fox family of transcription factors

Select the most specific GO term possible

Case 1: “nucleus” vs. “transcription factor complex”

```
GO:0005634 nucleus
GO:0005654 nucleoplasm
GO:0005667 transcription facto
```

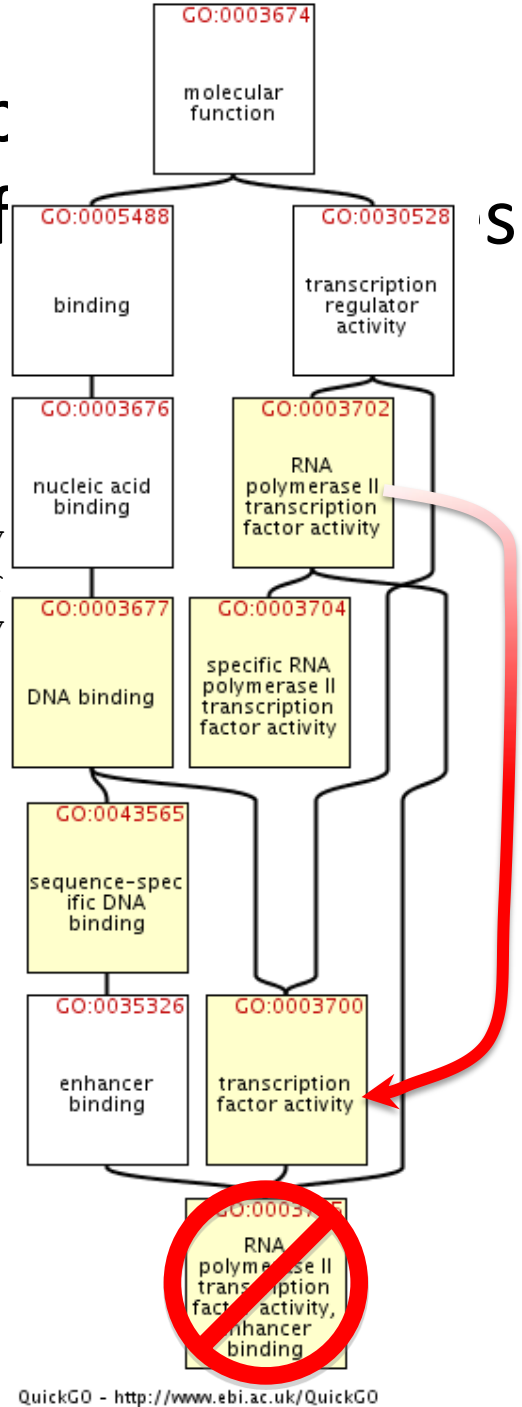


Select the most specific GO term possib
Case 2: DNA binding and transcription f s

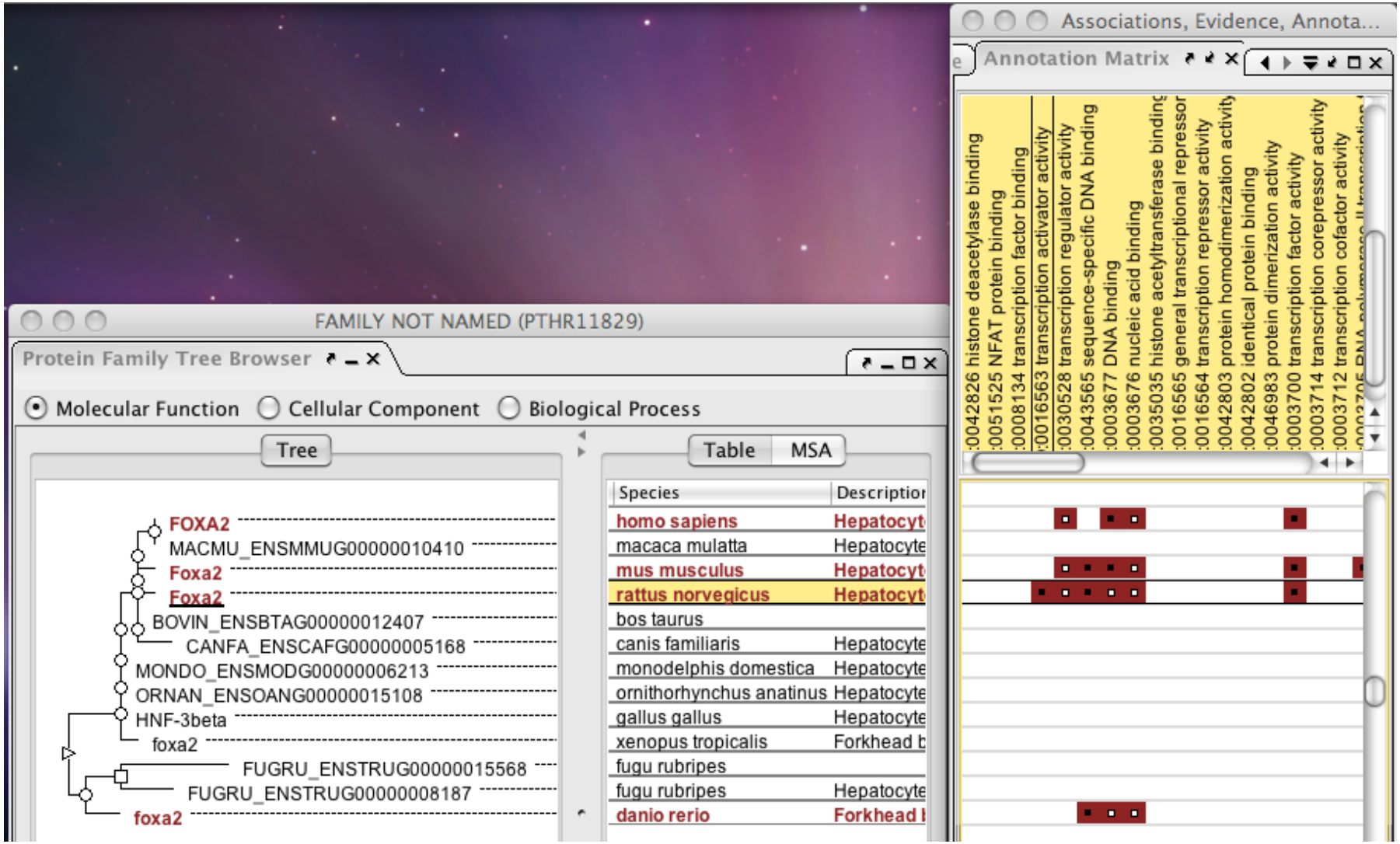
```
GO:0003677   DNA binding
GO:0043565   sequence-specific DNA binding
GO:0003700   transcription factor activity
GO:0003702   RNA polymerase II transcription factor activity
GO:0003704   specific RNA polymerase II transcription factor
GO:0003705   RNA polymerase II transcription factor activity
```

Shouldn't GO:0003702 "RNA polymerase II transcription factor activity" be a child of GO:0003700 "transcription factor activity" and not a direct child of GO:0030528 "transcription regulator activity"?

http://sourceforge.net/tracker/?func=detail&aid=704985&group_id=36855&atid=440764



Scope of annotations



Rule 2: Apply the term as widely as possible

- Some terms can be applied to all family members
  - “regulation of transcription factor activity”
  - “regulation of gene-specific transcription from RNA polymerase II promoter”
- Some terms can be applied to all metazoa
  - “tissue development”
  - “embryonic development”
  - “pattern specification process”
- More specific terms can be applied to specific clades
  - pos/neg regulation
  - *in utero* embryonic development

Annotation changes
Type 1: Small errors

"I can't find [FoxO3] mentioned in PMID 12930811, source of 2 CC annotations. Could there have been a typo?"

"It's a typo. J:88189 instead of J:88198. Fixed in MGI now."

http://sourceforge.net/tracker/?func=detail&aid=2997218&group_id=36855&atid=1040173

Annotation changes
Type 2: More specific term available

- Foxa1 & Foxa2: “branching morphogenesis of a tube” IGI
- Paper title: “Compensatory Roles of Foxa1 and Foxa2 during Lung Morphogenesis”
- GO:0060441 “branching involved in lung morphogenesis” added to GO after the paper was curated

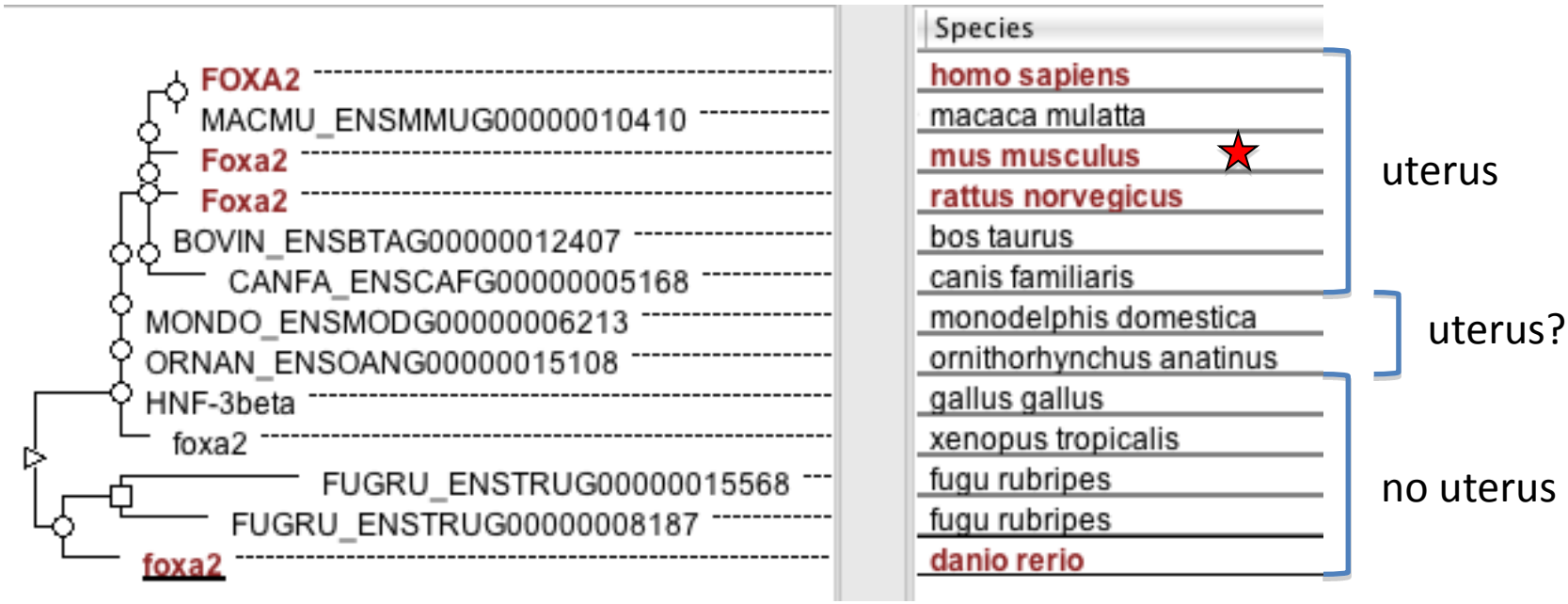
http://sourceforge.net/tracker/?func=detail&aid=2998353&group_id=36855&atid=1040173

Results (so far)

Annotation statistics			
PANTHER family	10046	22573	11829
Representative gene	Lon protease	HPRT	Forkhead
Proteins (RefGenome)	31	71	277
Proteins (all)	115	233	672
Up-propagations	20	46	387
Protein ISS annotations	1085	1900	14858
Annotation suggestions			
Accepted		1	10
Rejected			1
Pending	1	10	12
Ontology suggestions			
Accepted		3	2
Rejected			1
Pending			2

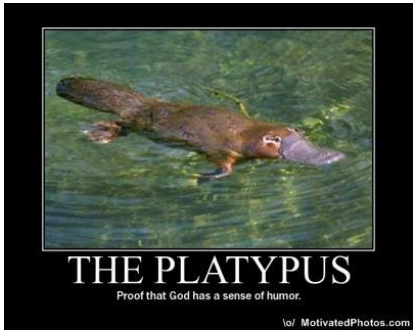
QuickTime™ and a
decompressor
are needed to see this picture.

Apply the term as widely as possible: "in utero embryonic development"



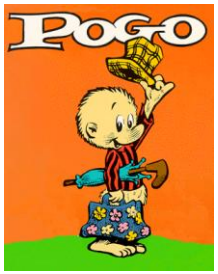
Monodelphis domestica

Gray Short-tailed Opossum (S. America)
(Marsupial)

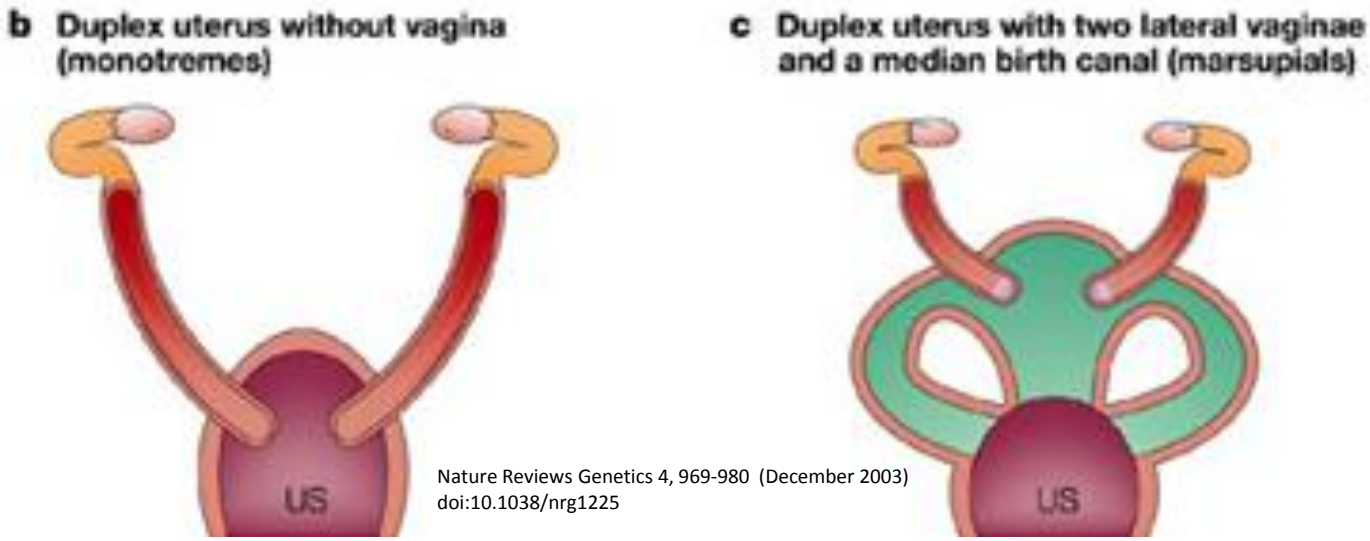


Ornithorhynchus anatinus

(Monotreme)

- Does the p ve a uterus?
- What abou 

sum?



Nature Reviews Genetics 4, 969-980 (December 2003)
doi:10.1038/nrg1225

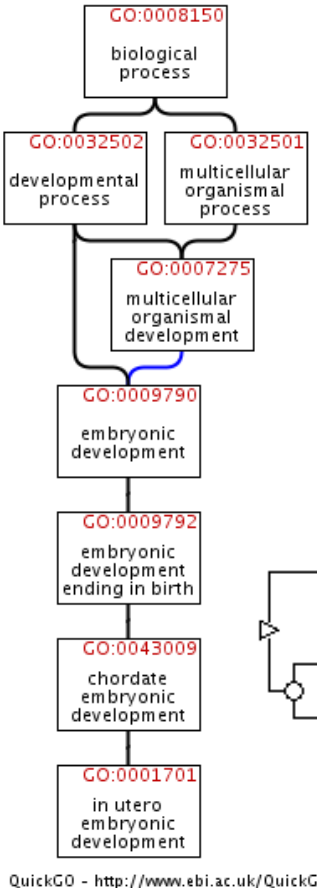
GO:0001701 “in utero embryonic development”: “The process whose specific outcome is the progression of the embryo in the uterus over time, from formation of the zygote in the oviduct, to birth.”

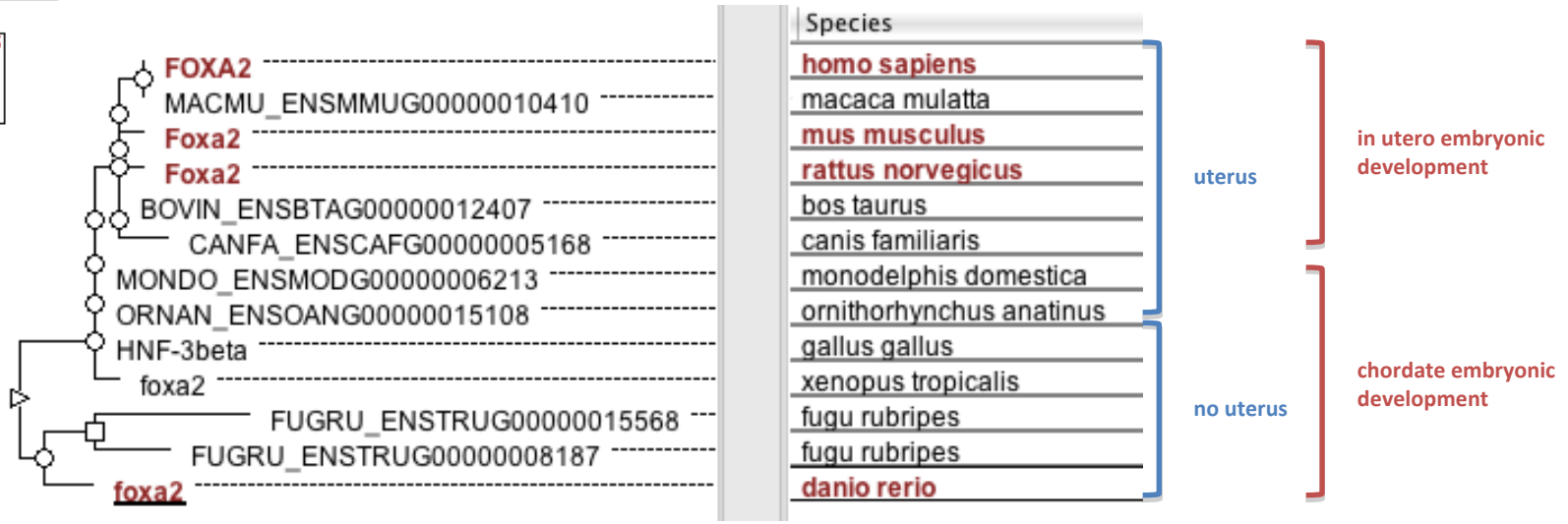
day 7 p.p.

http://www.public.iastate.edu/~zoogen/opposum.html

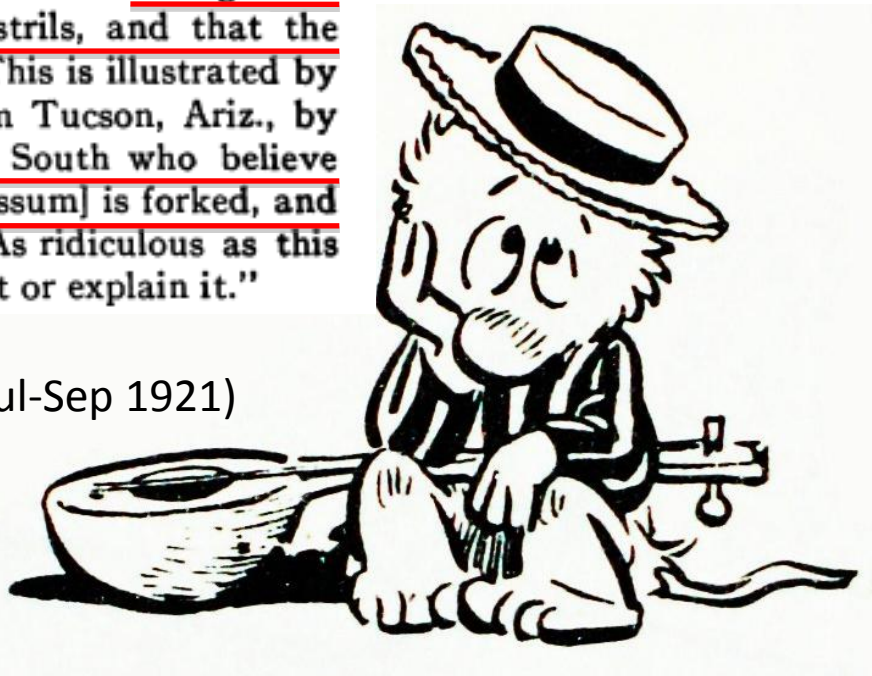
day 20 p.p.

Choose the most specific term.
Apply as broadly as possible.





Taxon constraints for GO terms?

TRADITIONAL BELIEF CONCERNING THE GENERATION OF THE OPOSSUM (*Didelphys virginiana*, L.). — The opossum figures in the folk-lore of the North American Indians, the negroes and whites of the South, and the natives of Mexico. Such stories as the writer has seen concern the opossum's hairless tail, its broad grin when in the attitude of fear or anger, its habit of feigning death ("playing 'possum"), and its use for food. There seems to be no published reference to the prevalent notion regarding the mode of generation in this animal, — a notion which is deeply ingrained in the popular mind, from New Jersey to the Rio Grande River. This general belief is that the opossum copulates through the nostrils, and that the female blows the fruit of conception into the pouch. This is illustrated by the following note, dated April, 1921, sent to me from Tucson, Ariz., by J. Ross Thurman: "There are many old men in the South who believe the old Southern adage that the organ of the male [opossum] is forked, and the intercourse takes place in the nose of the female. As ridiculous as this sounds, I have never heard of any one who could deny it or explain it."

The Journal of American Folklore 34(133):321-3 (Jul-Sep 1921)
http://www.jstor.org/stable/535156