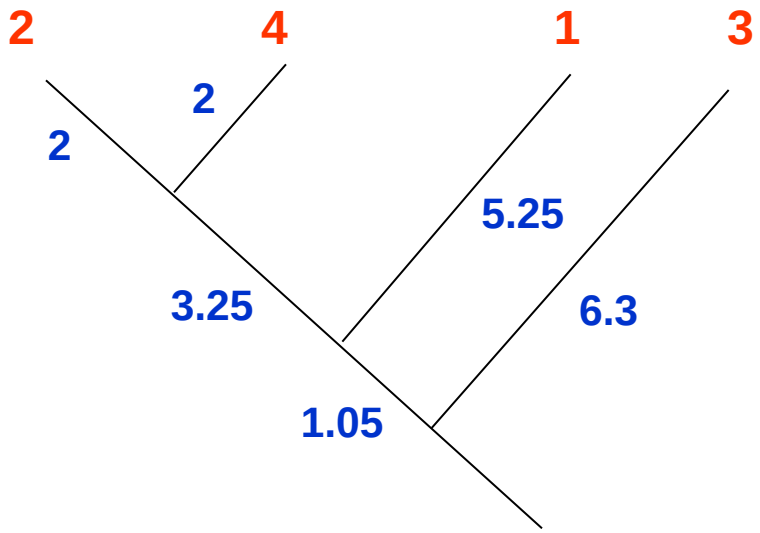
Practice question #1a

UPGMA method

	1	2	3
2	10		
3	11	14	
4	11	4	13

Gap scored as difference

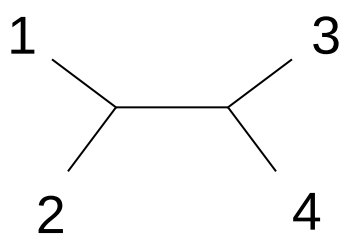


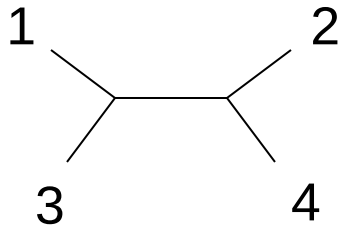
Practice question #1b

Maximum parsimony method

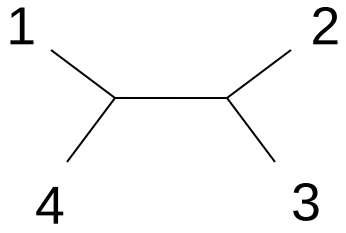
```
1    G T C T T T C G T C A G G - T G G A C G T C C C C A G C C C A G
2    G T C T C G A G C C A G T C T G G G T G T T C T C A G C C C A G
3    G T C T T T G A - C A A G A T G A T G A T C C T T A G C C C A G
4    G T C T C G T A A C A G T C T G G G - G T T C T C A G C C C A G
```

Informative sites: positions 5, 6, 8, 13, and 22





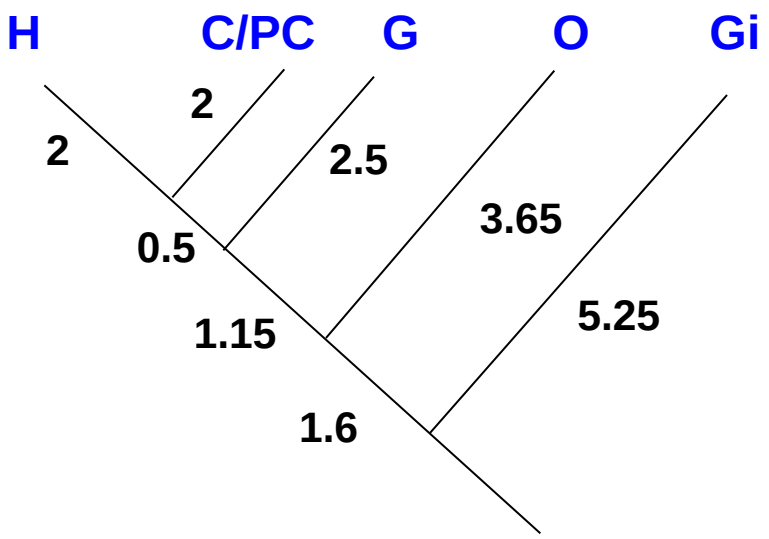
*



Total number of dots	**9**	**6**	**10**
		Preferred tree	

Practice question #2

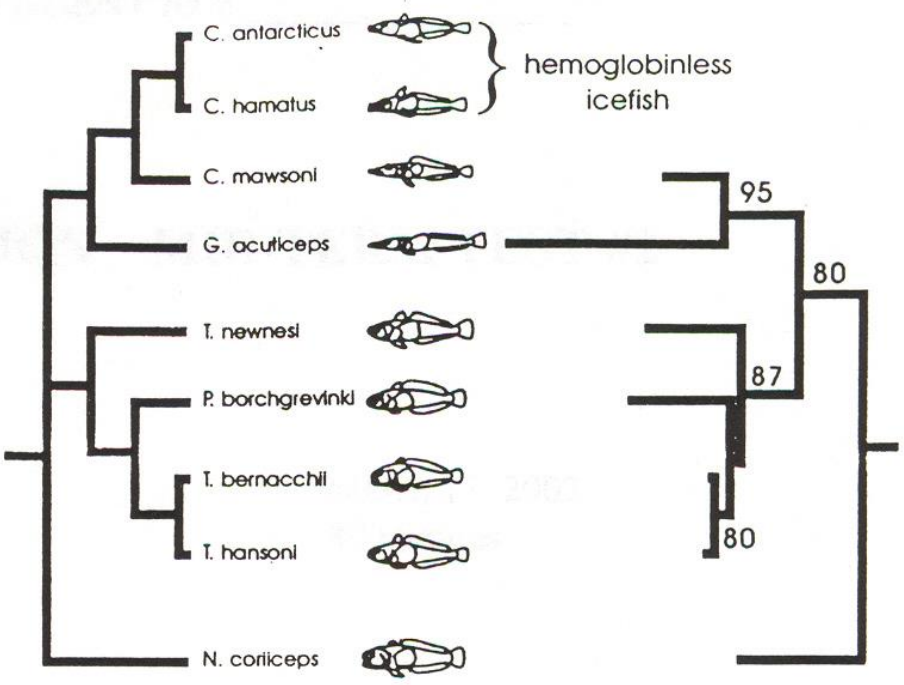
	H	C/PC	G	O
C/PC	4			
G	5	5		
O	9	5	8	
Gi	11	9	10	12



Note: Score as difference if restriction site is not at exactly same position or if different symbol at same position

Practice question #3

The figure says these fish lack hemoglobin, so are not represented on the gene tree



www.icefish.neu.edu/ photo/

"Are red blood cells really necessary? ... a fish species that lives without them"

Species tree

Shows the known phylogenetic relationships among a set of organisms

Often reconstructed using molecular data (as well as other) - eg. rRNA genes, histone genes…

Gene tree

(using hemoglobin sequences) because interested in its evolutionary behaviour in these fish

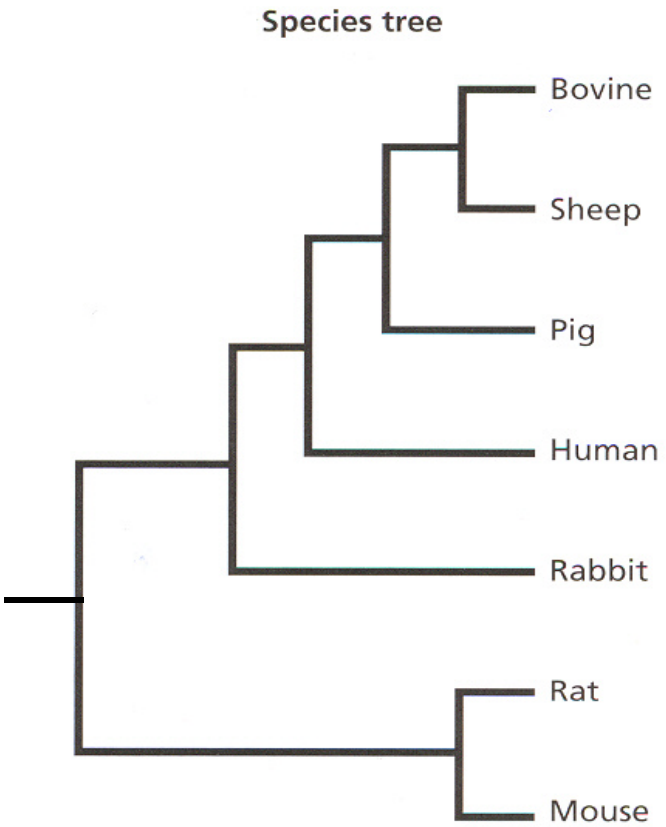
Do the 2 trees show the same topology?

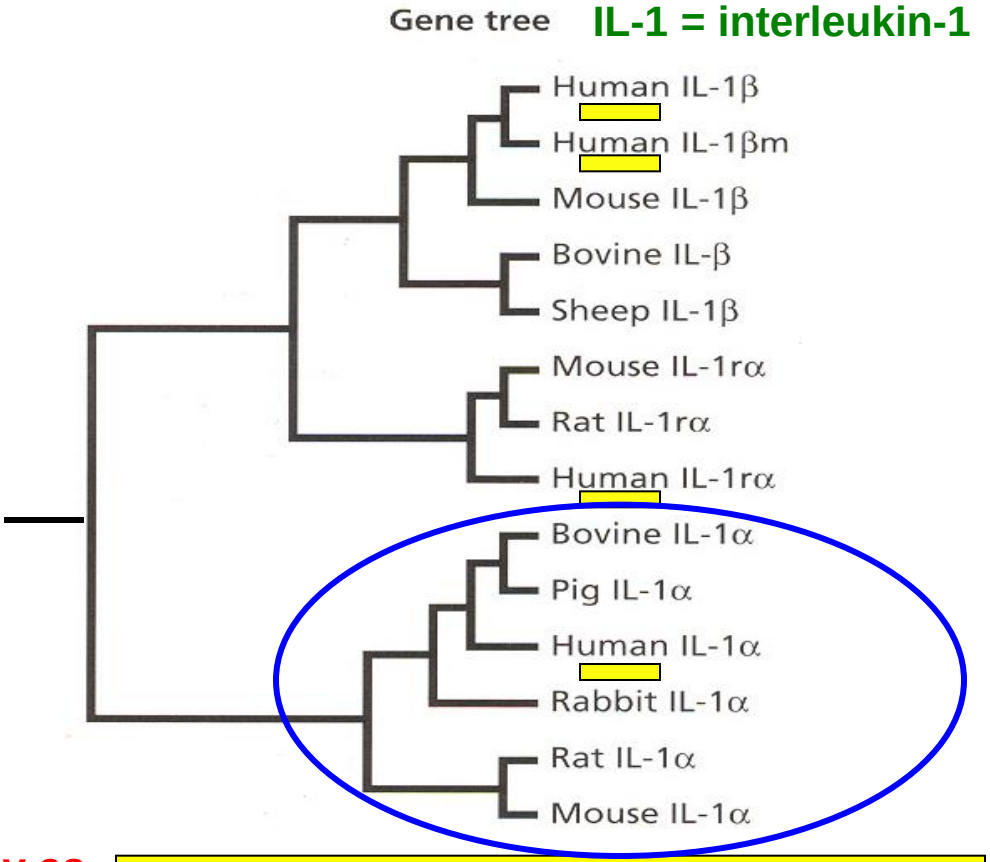
Implications of differences in branch lengths on gene tree?

Long branches: reduced functional constraint? or evolving new function (adaptive evolution)…?

Bargelloni *PNAS* 95:8670, 1998

Practice question #4





Does gene tree show the same topology as the species tree? (eg. for IL-1α clade)

Clade has same topology as species tree, so IL-1α type gene present in common ancestor & accumulating nt subs in predicted manner

How many copies of IL-1 gene?

- 4 in human genome so at least 3 duplication events – predict two in common ancestor & one in human-specific lineage

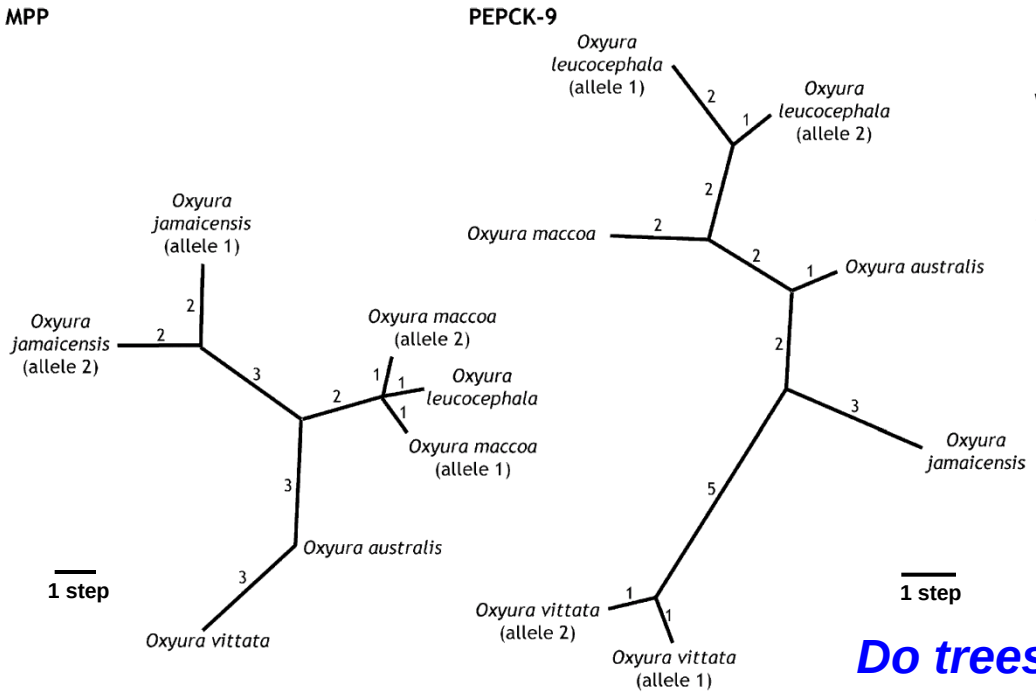
When did duplication events occur?

Certain copies of genes lost from specific lineages vs. incomplete data sets for some organisms?

sheep lacking IL-1α gene or incomplete data set

Practice question #5

- 2 data sets: MPP & PEPCK introns analyzed using maximum parsimony



numbers: # steps

Within a tree, short branch lengths reflect fewer steps, so shorter time (assuming molecular clock hypothesis)

... ie. ducks separated by short branches are more closely--related

Do trees have same topology?

Why do branch lengths differ between trees?

- **faster rate of nt sub in PEPCK-9 intron?**
- **or slower rate in MPP intron?**

maybe MPP constrained because contains important info (eg. cis-elements for regulation)

- **or maybe PEPCK-9 intron is simply longer than MPP intron, ?**

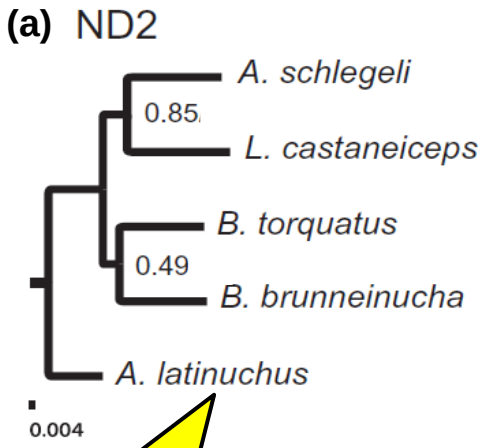
so there are more informative sites...

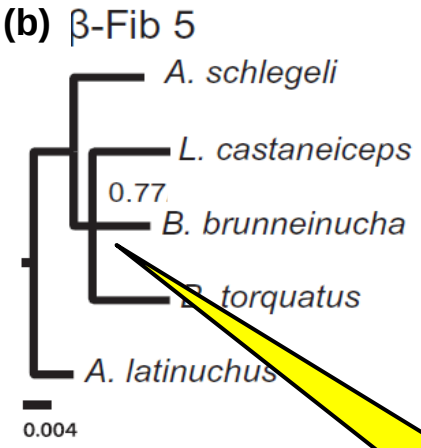
Practice question #6

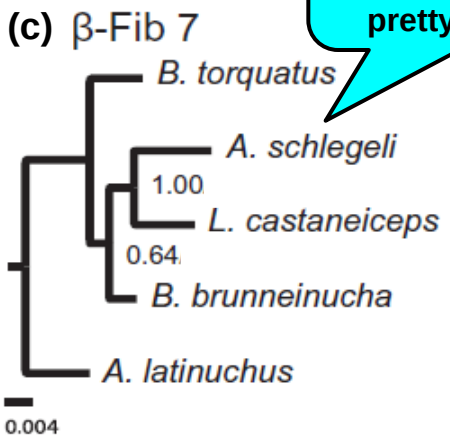
(ie. they want a species tree)

- interested in phylogenetic relationships among these birds

- but trees using these 3 genes showed different topologies







Clade looks solid (good bootstrap value & ND2 one is pretty good too)

Researchers chose this bird as outgroup (they knew it was less-closely related) - as is done for relative rate tests

Trichotomy with only "moderate" bootstrap support

Should collapse (polytomy with no distinction among the 4 species)

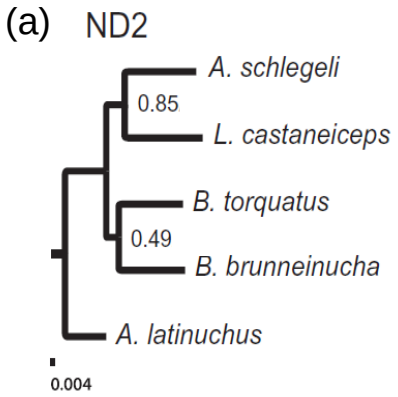
So very little info from Fib5 gene (maybe too few informative characters…)

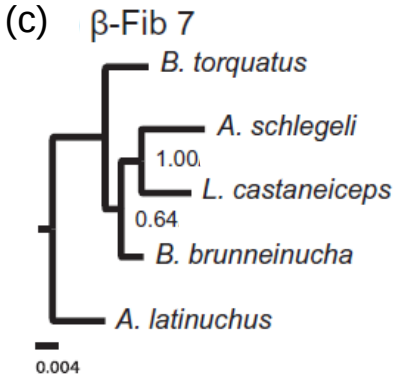
- possible reasons for different topologies between (a) and (c) for the two Buarremon species ?…

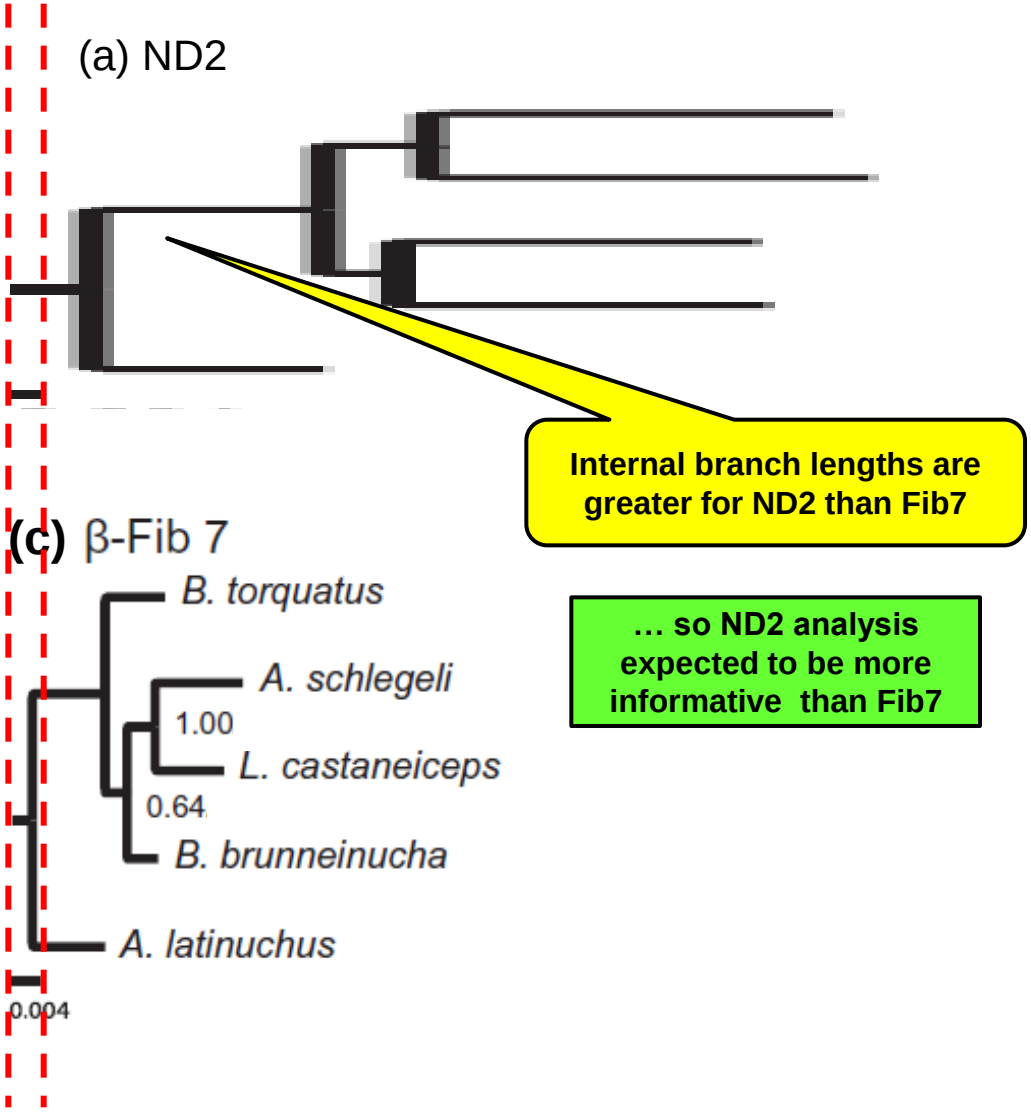
- confidence for this topology is low (cf. bootstrap values)

… and scales of trees are very different !!

(see next slide where I've "stretched" tree (a) to same scale as tree (c)



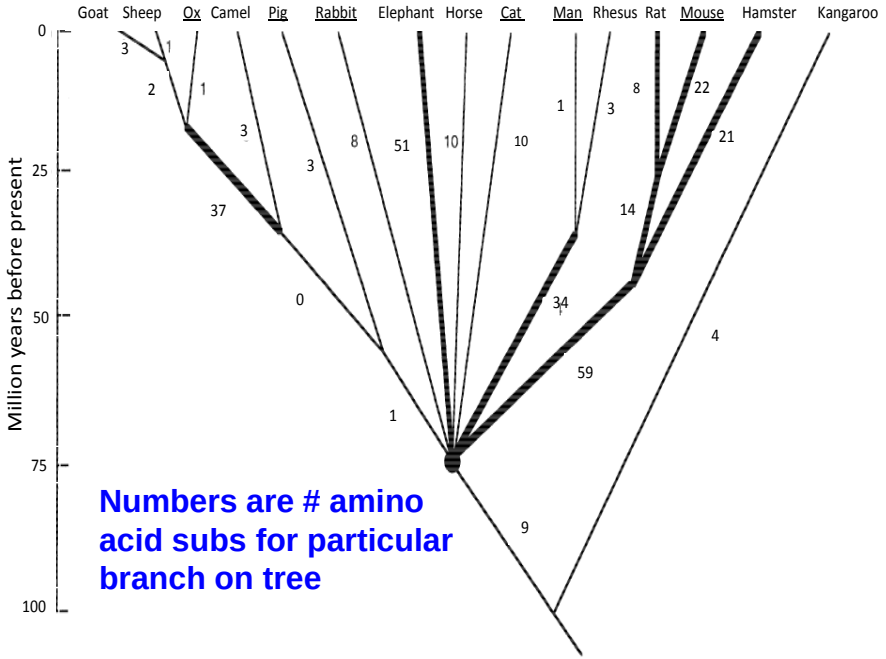




Cautionary note: high bootstrap values give measure of consistency, but they don't guarantee "true" topology

Practice question #7 (this was Question 6 in the first set)

	#1 k_A	#2 k_S
Human	0.111	1.33
Ox	0.131	1.01
Pig	0.040	1.28
Cat	0.052	1.60
Rabbit	0.050	1.21
Mouse	0.261	1.58



- pairwise comparisons of nt sequences with outgroup(s), score differences
- alignment, gap penalty, determine # syn vs. non-syn sites, normalize for length…

Molecular clock hypothesis

… and punctuated (or episodic) evolution

- know that rodents have high sub rate (for syn & non-syn sites) - high metabolic rate…

But unequal # aa subs for prolactin since time that rat & mouse shared common ancestor

But elephants have high number of aa subs too (large animal, not high metabolic rate…)

- reduced functional constraint ? (“relaxation of purifying selection”)

or adaptive evolution?