

The bioGENESIS-ATBC2012 Training
workshop "Evolutionary Approaches to
Biodiversity Science"
16-18 June 2012, Bonito, Brazil

“Phylogenetic and functional diversity
(including PD) and phylogenetic
conservation planning”

Dan Faith
The Australian Museum

- Cam Webb

- Phylogenetic community structure and Beta Diversity measures

- Dan Faith

- Phylogenetic and functional diversity (including PD) and phylogenetic conservation planning

- Félix Forest

- Biodiversity informatics (including data mining) and biodiversity measures (EDGE, PE)



bioGENESIS

Providing an evolutionary framework for biodiversity science



bioGENESIS Science Plan and Implementation Strategy



BioGENESIS

Providing an evolutionary framework for biodiversity science

available on the DIVERSITAS web site

Why conserve phylogenetic diversity?

- Think of phylogeny or the tree of life as a kind of heritage
 - Evolutionary heritage
- Phylogenetic diversity is an important target for biodiversity conservation because it represents current and future benefits for humans (evolutionary or “ecosystem services”).

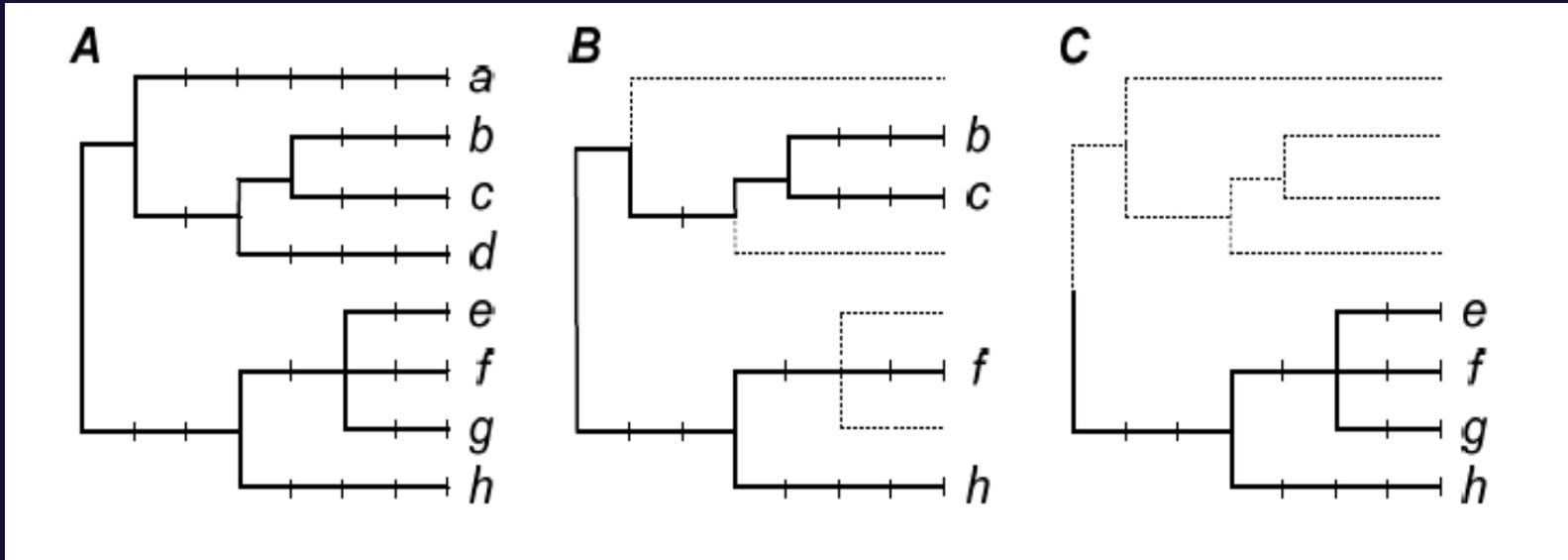
Faith et al (2010) Ecosystem services: an evolutionary perspective on the links between biodiversity and human well-being. *COSUST*

How does phylogeny provide a measure of conservation value?

- A phylogenetic diversity measure (PD; Faith 1992) helps quantify these current and potential future benefits ("option values") from the tree of life.
- PD links future benefits to estimates of "feature diversity".
- Phylogenetic patterns of evolutionary diversification predict feature diversity of sets of species.
- The total PD of a given set of species is the total phylogenetic branch length spanned (represented) by its member species.

PD – phylogenetic diversity Faith 1992

PD of a set of taxa = length of spanning path of the set on the phylogeny
(how much of the tree travelled over if connect up those taxa on the tree)



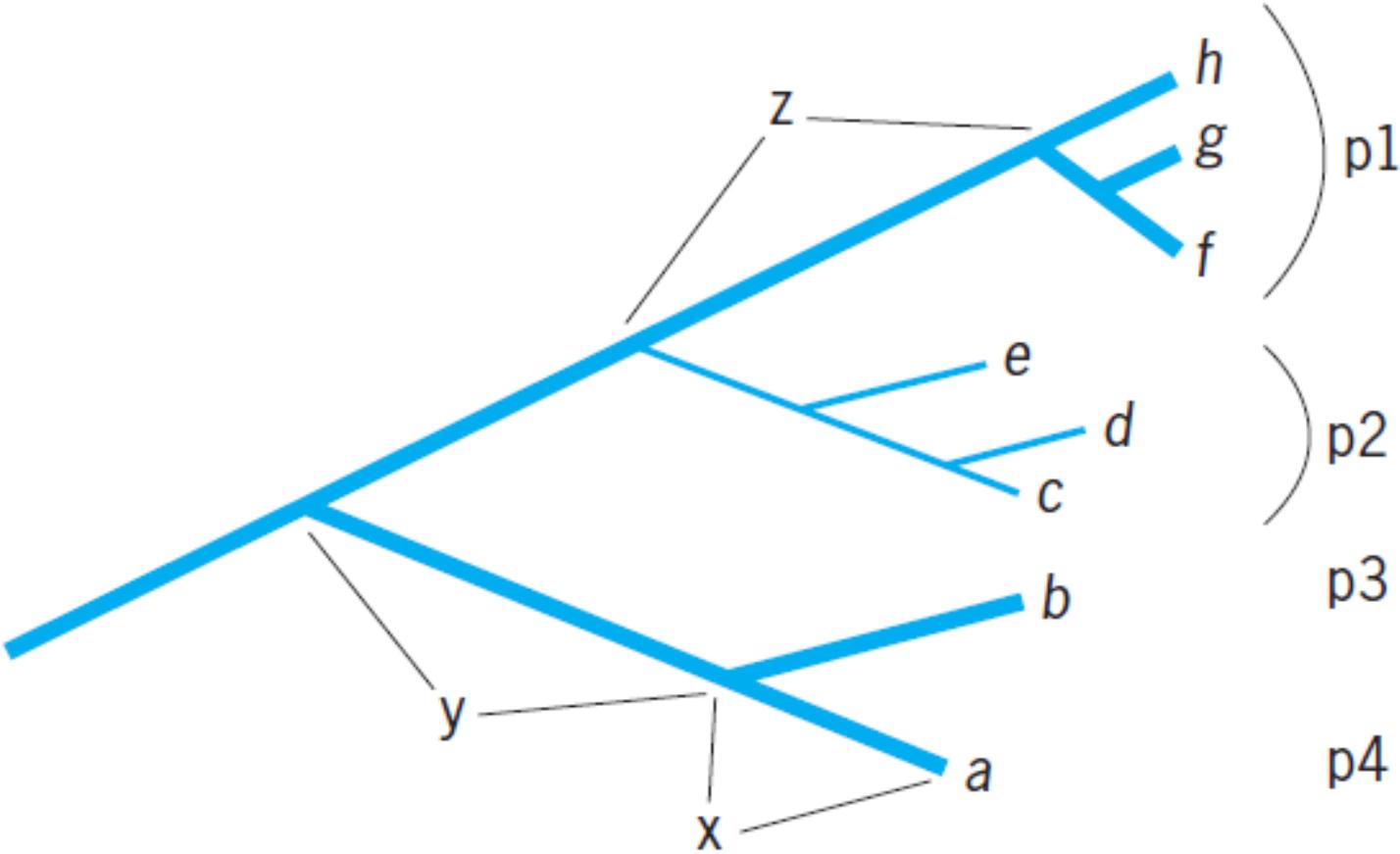
PD measures “feature diversity”

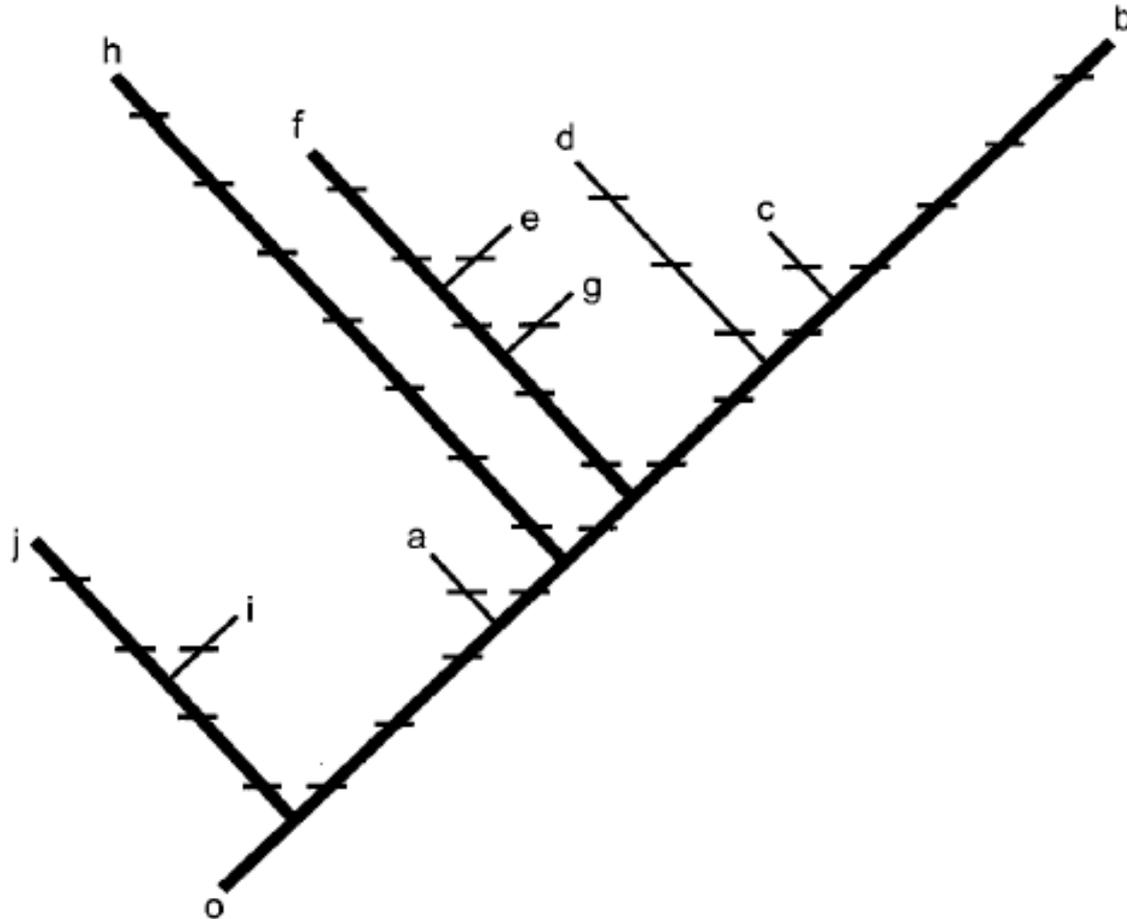
scenario *B* represents more feature diversity

Faith DP. *Biological Conservation* (1992).

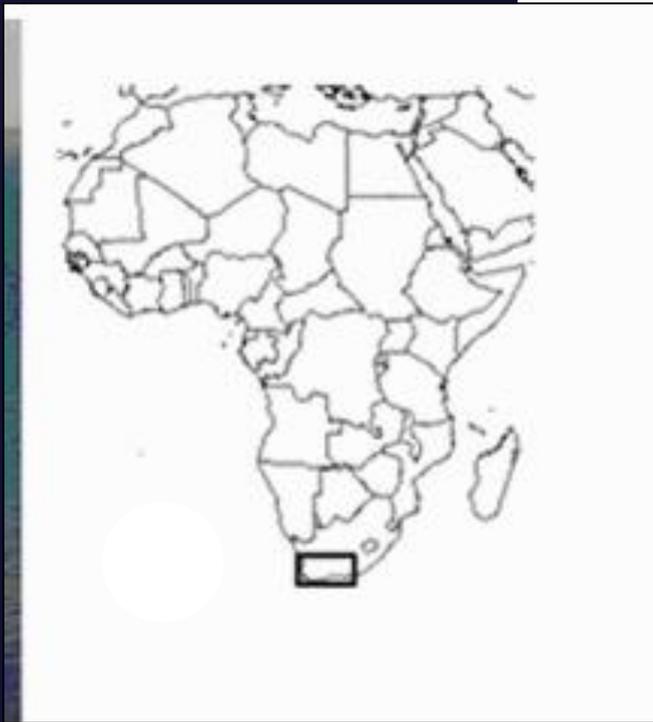
Faith DP. *Cladistics* (1992) 8:361-373.

Branches do not have to be time or clock-like

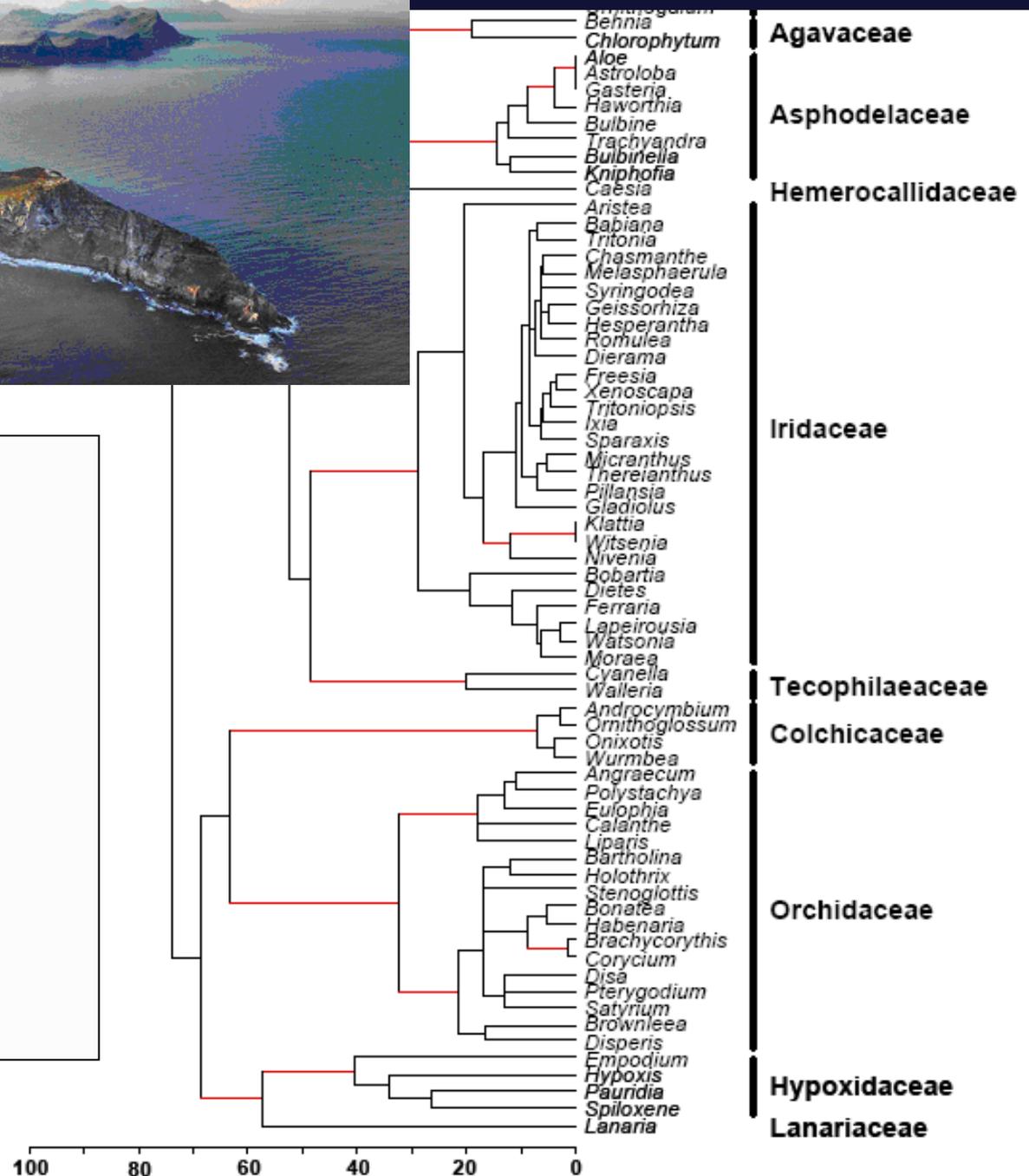




We use observed feature (characters) to infer the tree and branch lengths, and then use this phylogeny to make inferences about general feature diversity patterns among these taxa

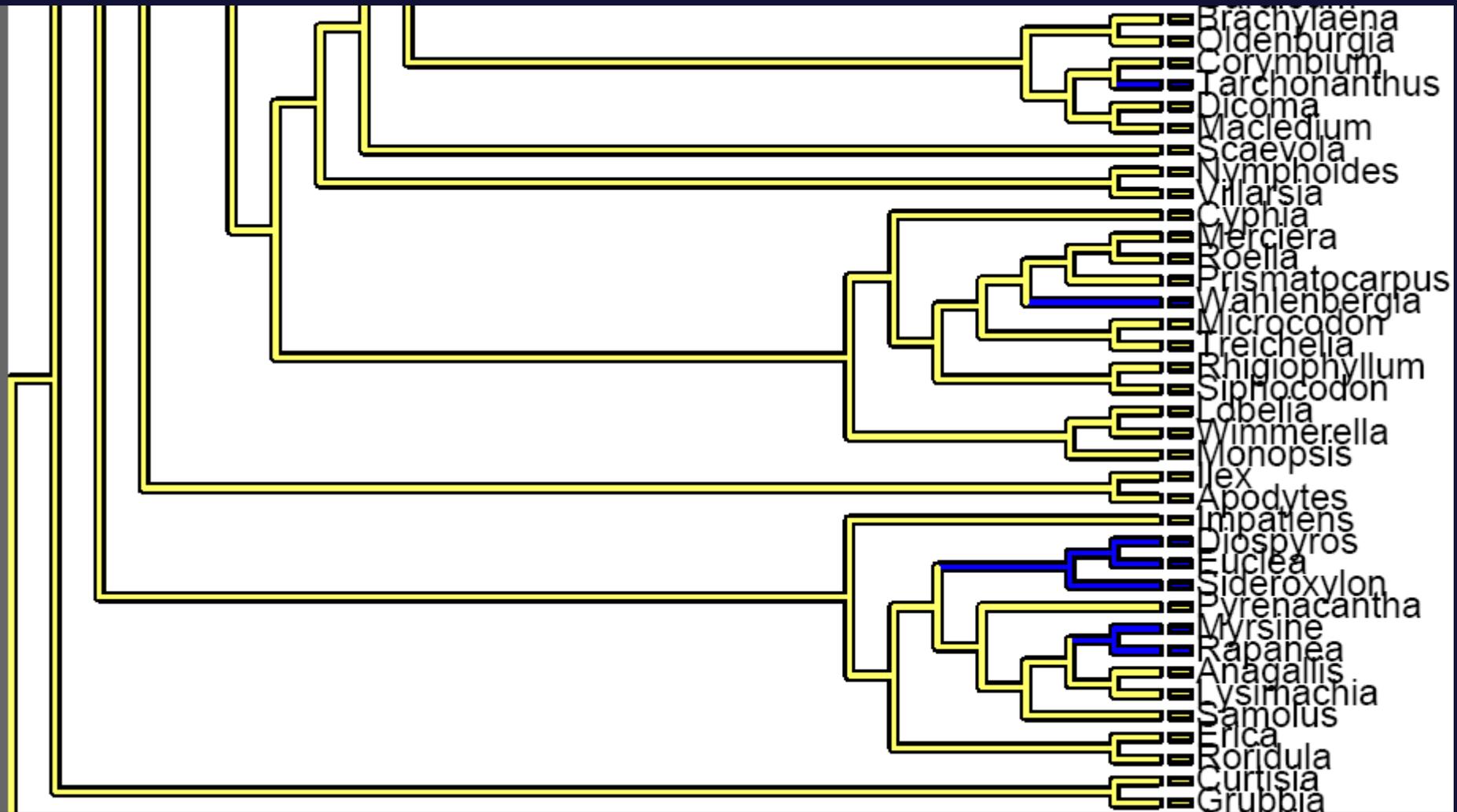


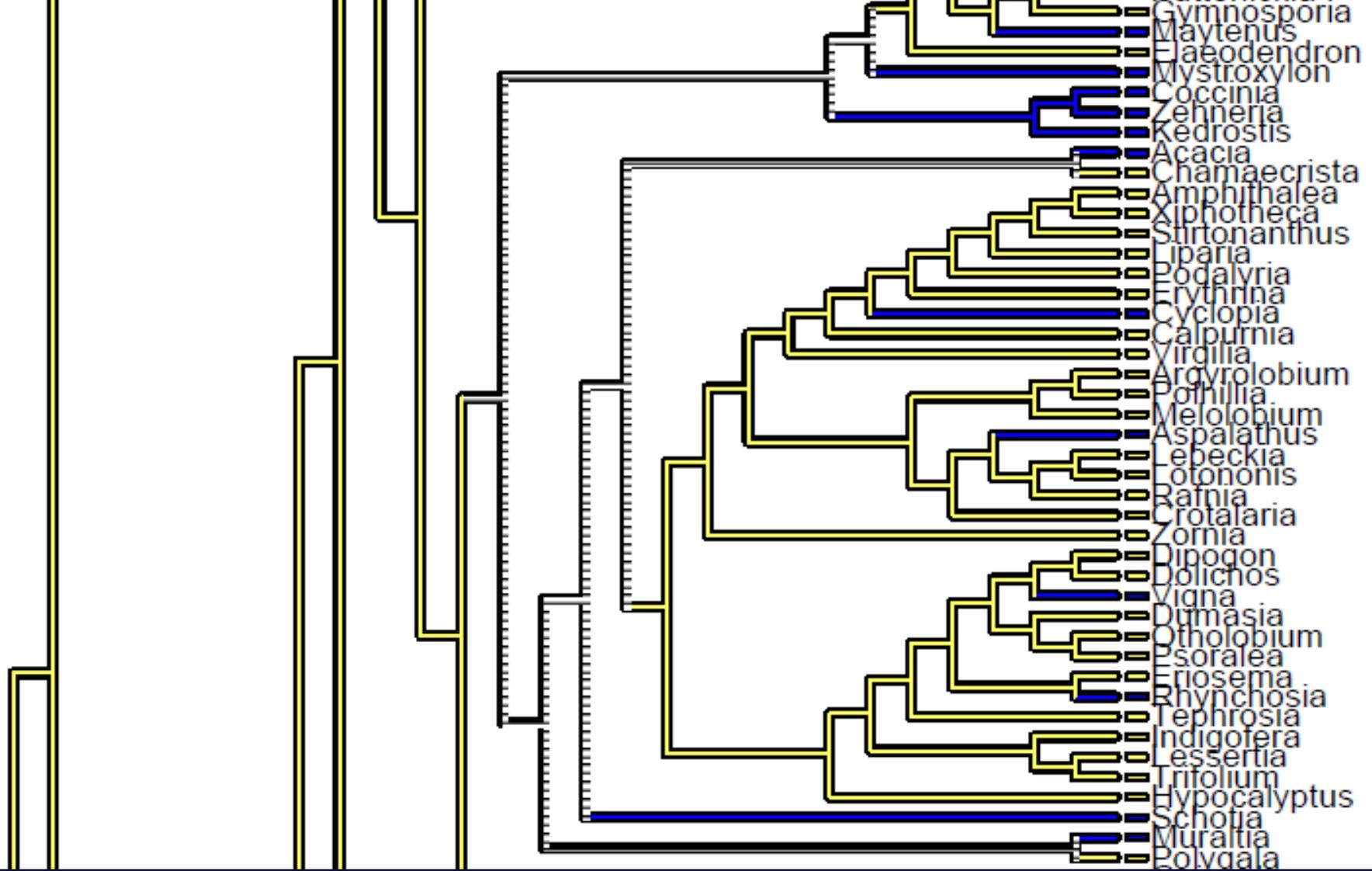
Forest et al. (2007)
Nature



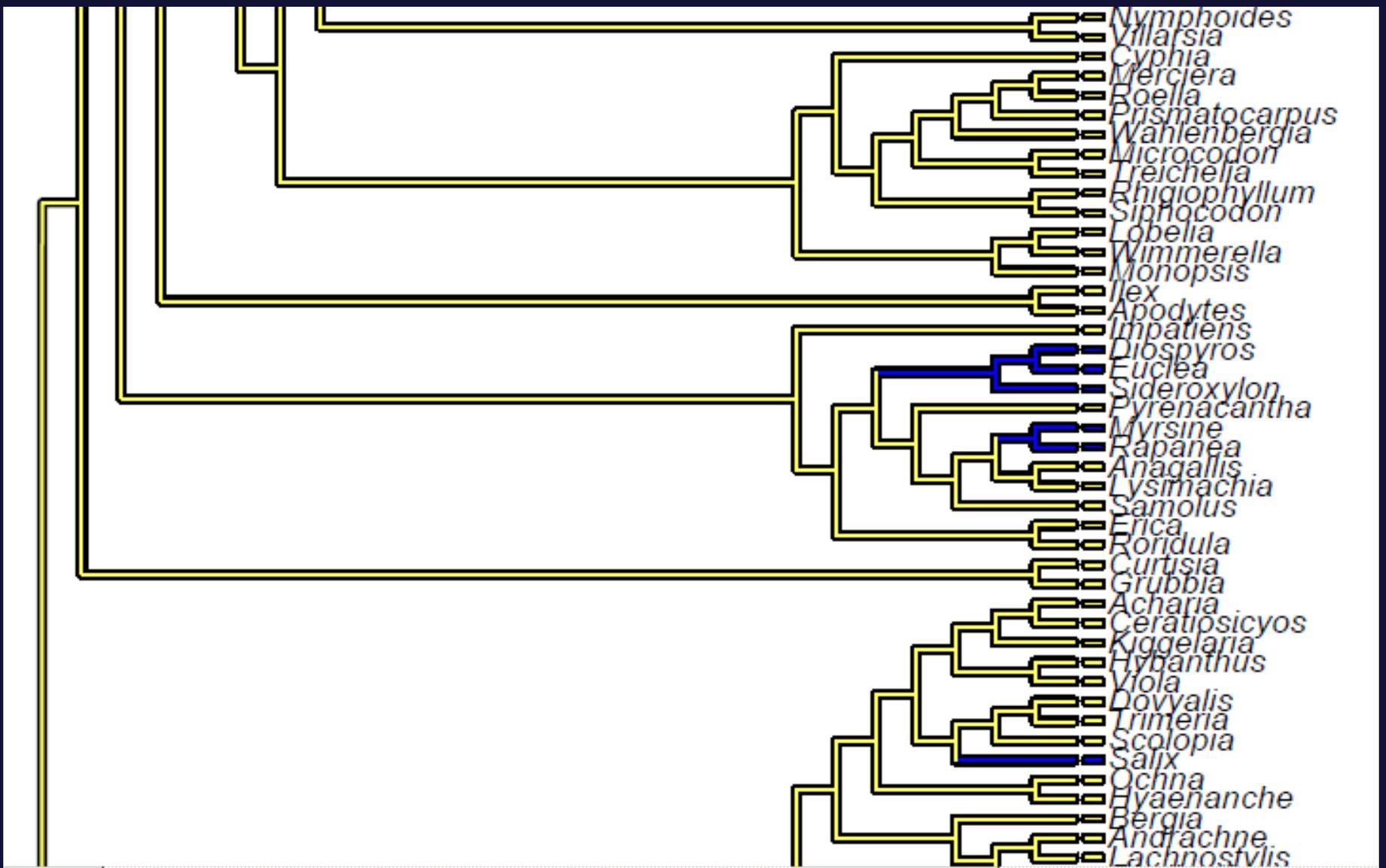
blue = genera in the Cape having species
of medicinal or economic importance

(as recorded in Survey of Economic Plants for Arid and Semi-Arid Lands)





Food uses



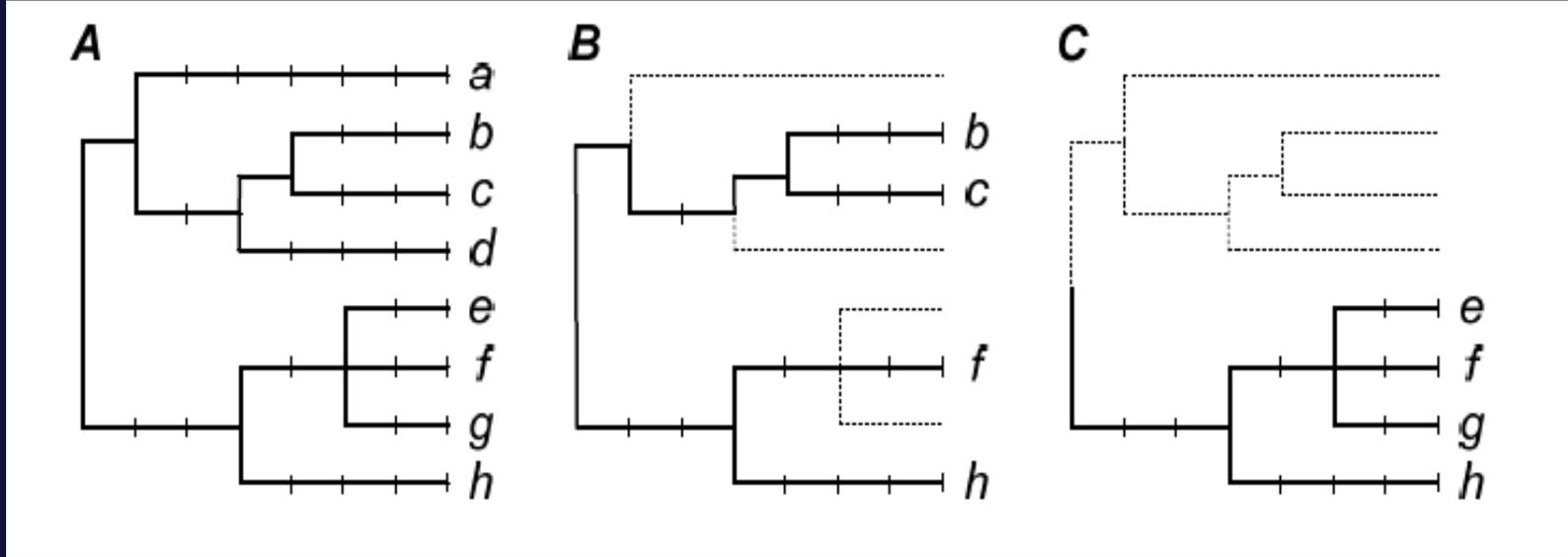
Medicinal uses

PD the best general predictor over different features subsets, so it best captures options for the future

PD property of “counting up” features has other useful implications

PD – phylogenetic diversity

Faith 1992



Phylogenetic ecology - take any conventional species-level index and re-express as a PD-based measure

Richness = total PD

Expected diversity = expected PD

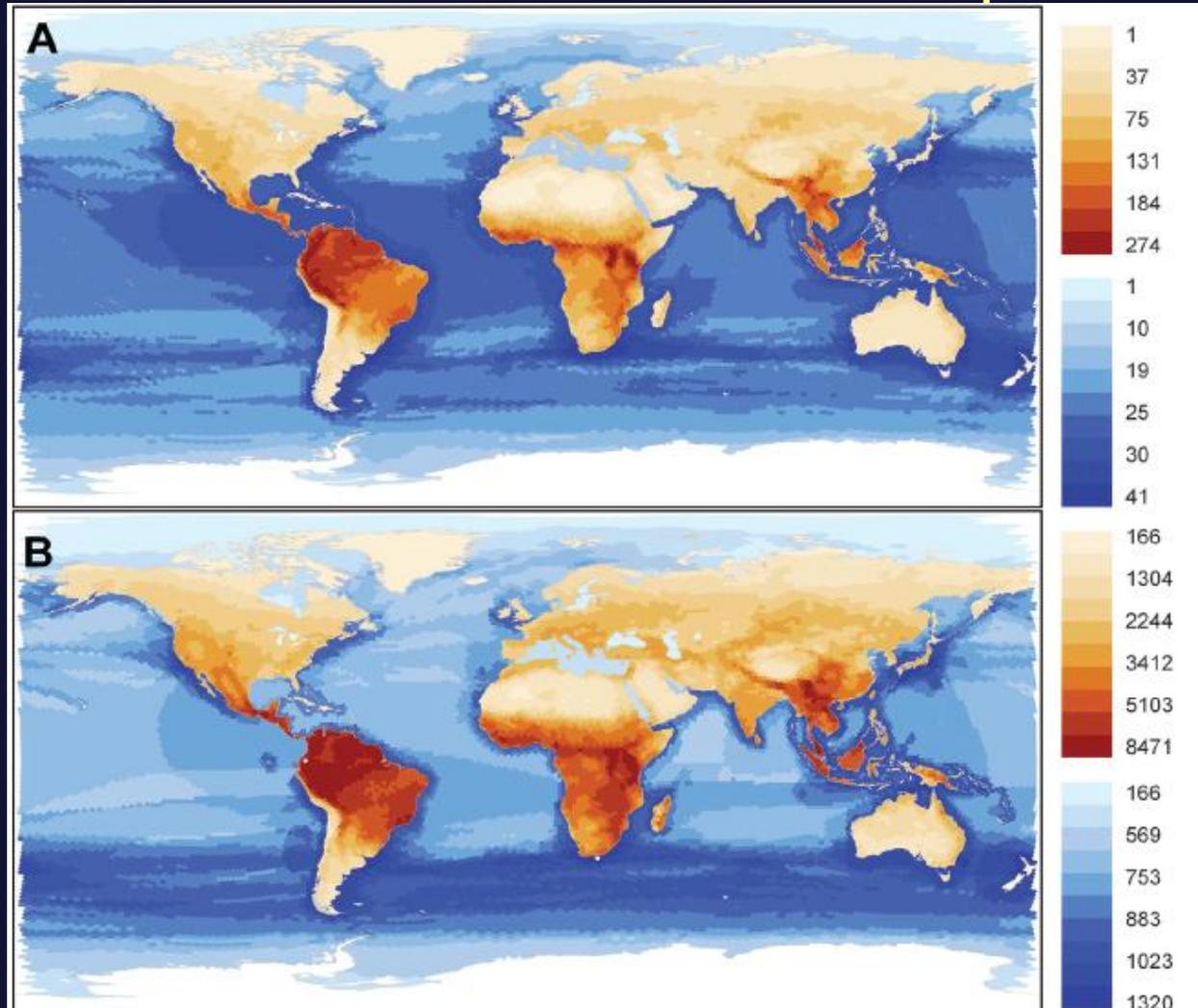
PD-Complementarity (gains & losses)

PD-Endemism (e.g. Faith et al 2004; Faith 1994)

PD-Dissimilarity between communities

PD analogues of Shannon-Weiner index and Simpson's index

Total PD often looks the same as total species diversity

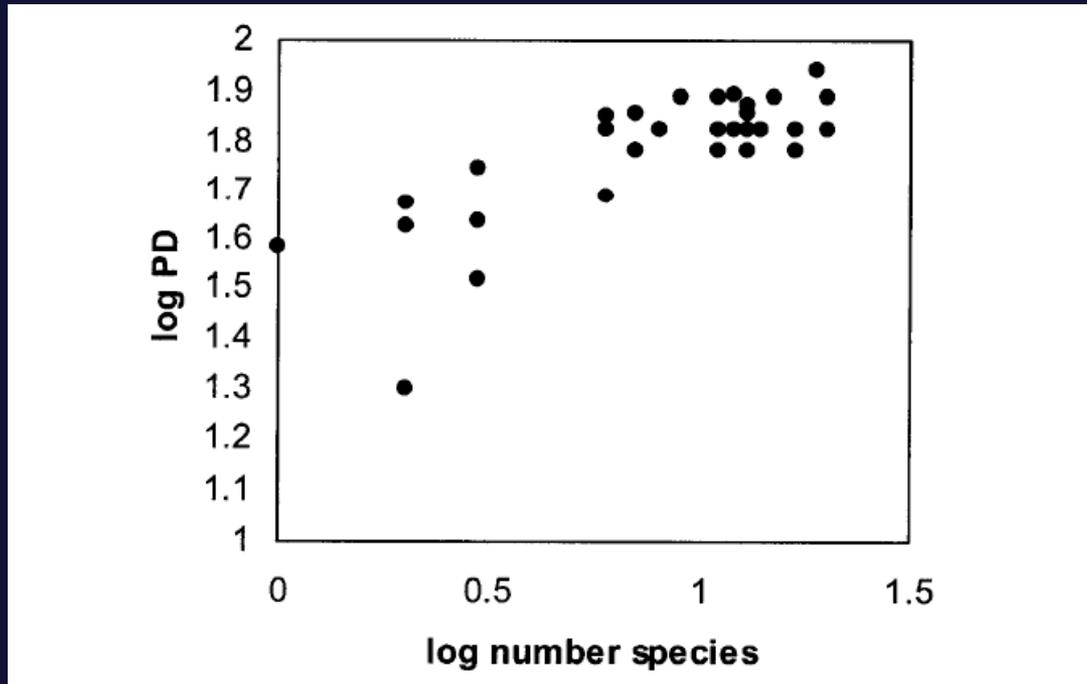


Schipper, et al. *Science* 322, 225 (2008)

Lots of PD applications. But popularity also means that term “PD” sometimes used for other methods, or as abbreviation for “phylogenetic diversity” generally

The PD – species relationship

- When number of species sampled is plotted against the PD value of the set, PD defines a species–phylogenetic diversity curve - analogous to species–area curve
- e.g. for data in Pilon et al (2006). Random taxon samples of different sizes from phylogenetic tree produce a roughly linear relationship in log–log space. (Faith & Williams, 2006; Faith 2008)



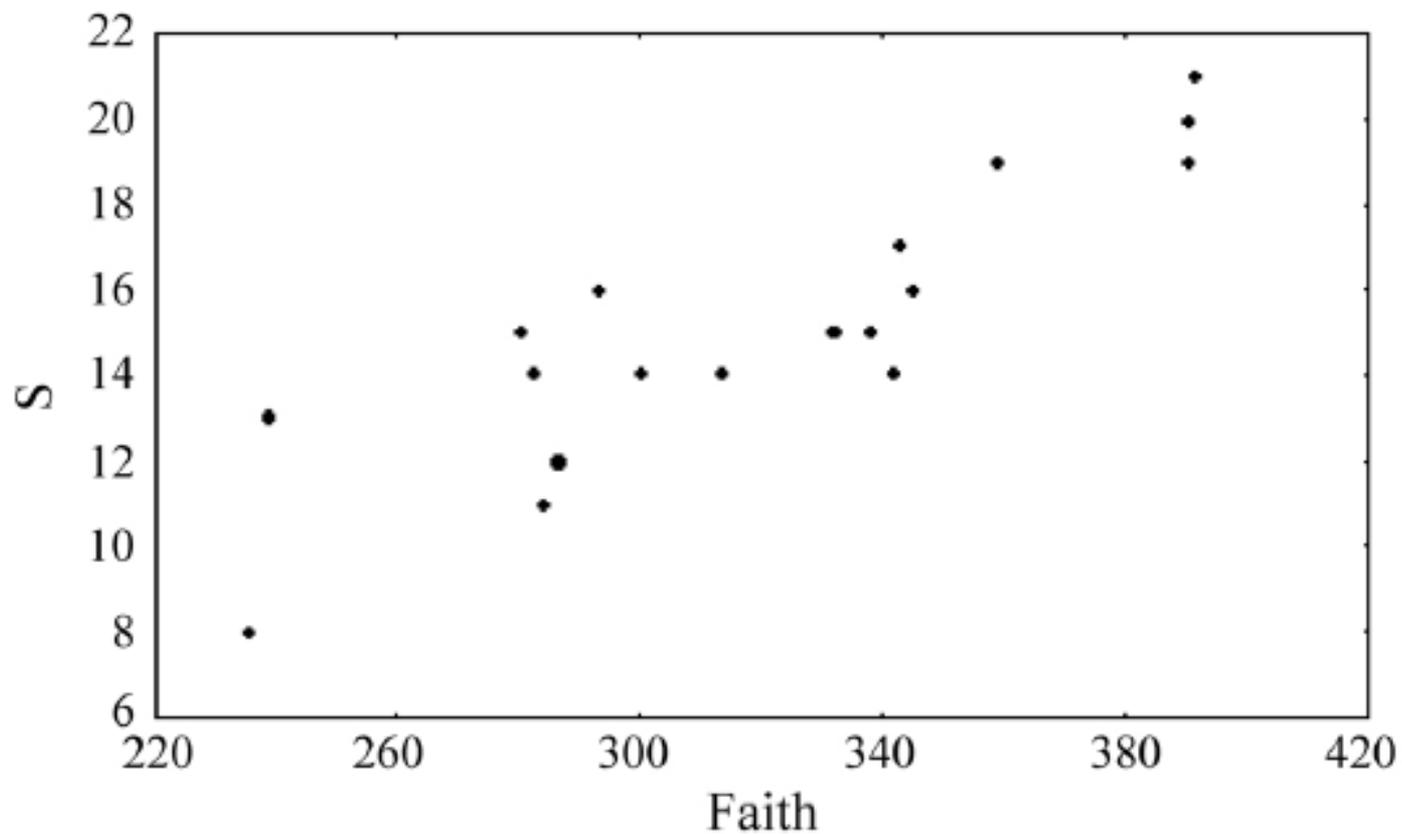


Figure 4 - Relationship between Faith's phylogenetic diversity index and the species richness, for the 18 conservation units.

e.g. Tôrres and Diniz-Filho (2004)

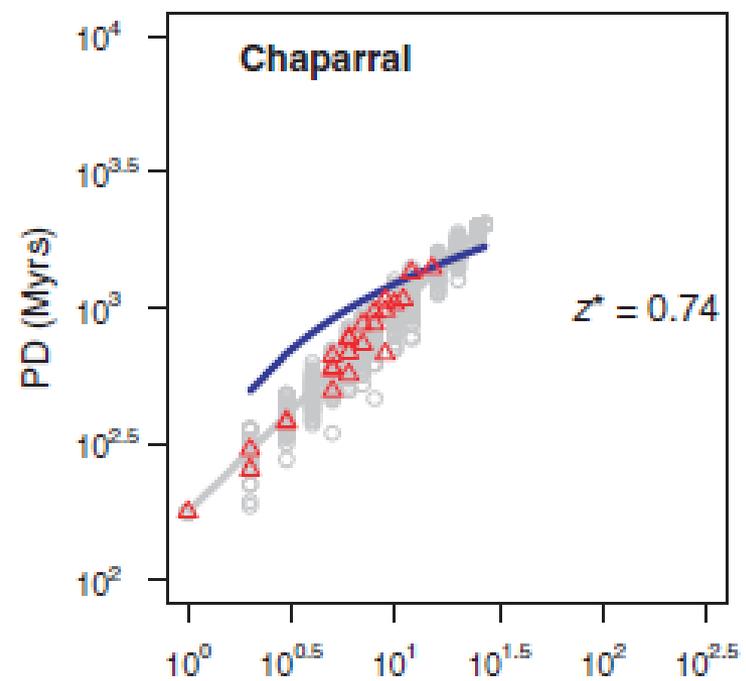
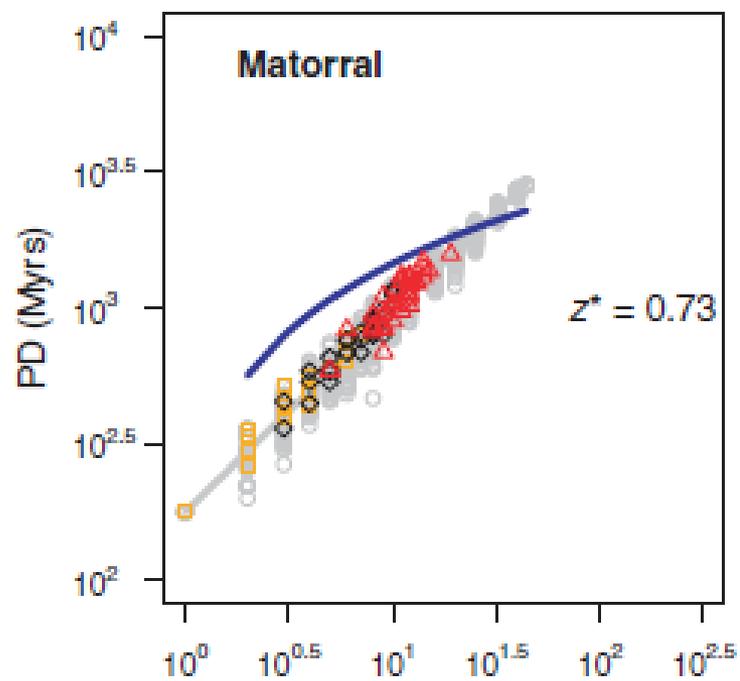
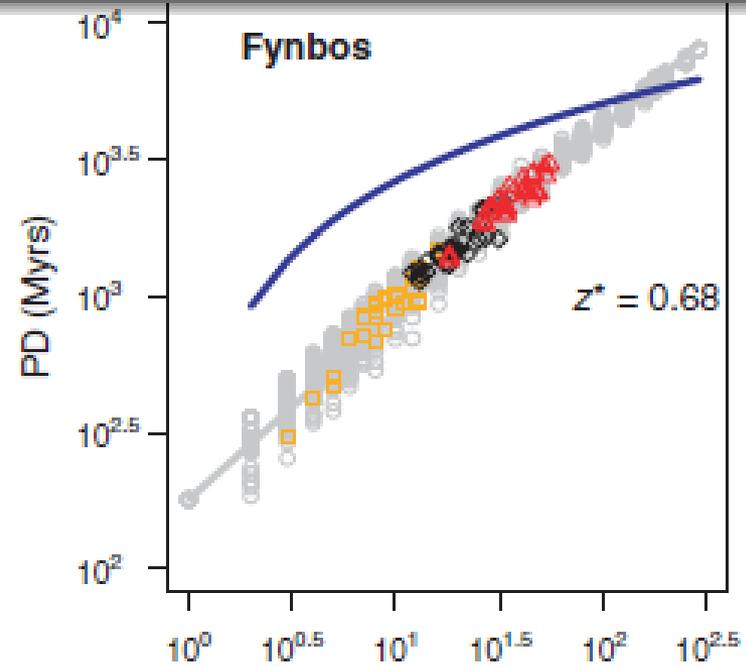
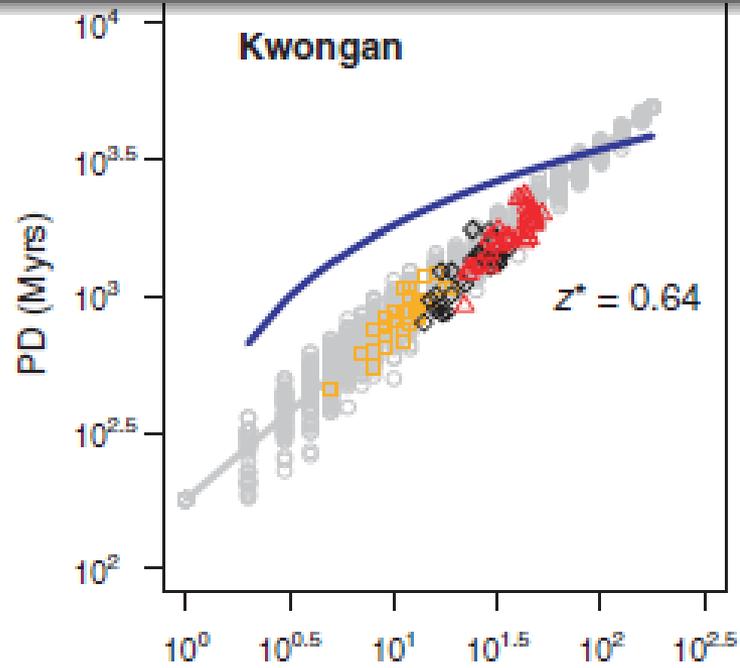
Phylogenetic autocorrelation and evolutionary diversity of Carnivora (Mammalia) in Conservation Units of the New World. *Genetics and Molecular Biology*, 27, 4, 511-516

Morlon et al (2011) found empirical support for this proposed power law model.

PD curves for 4 phylogenetic trees from 4 Mediterranean-type ecosystems.

For each species richness (S) randomly sample S species and calculate PD (do this 100 times).

Species-PD relationship well fit by a power law for all four phylogenies.



This relationship may be quite general

Irritable bowel syndrome

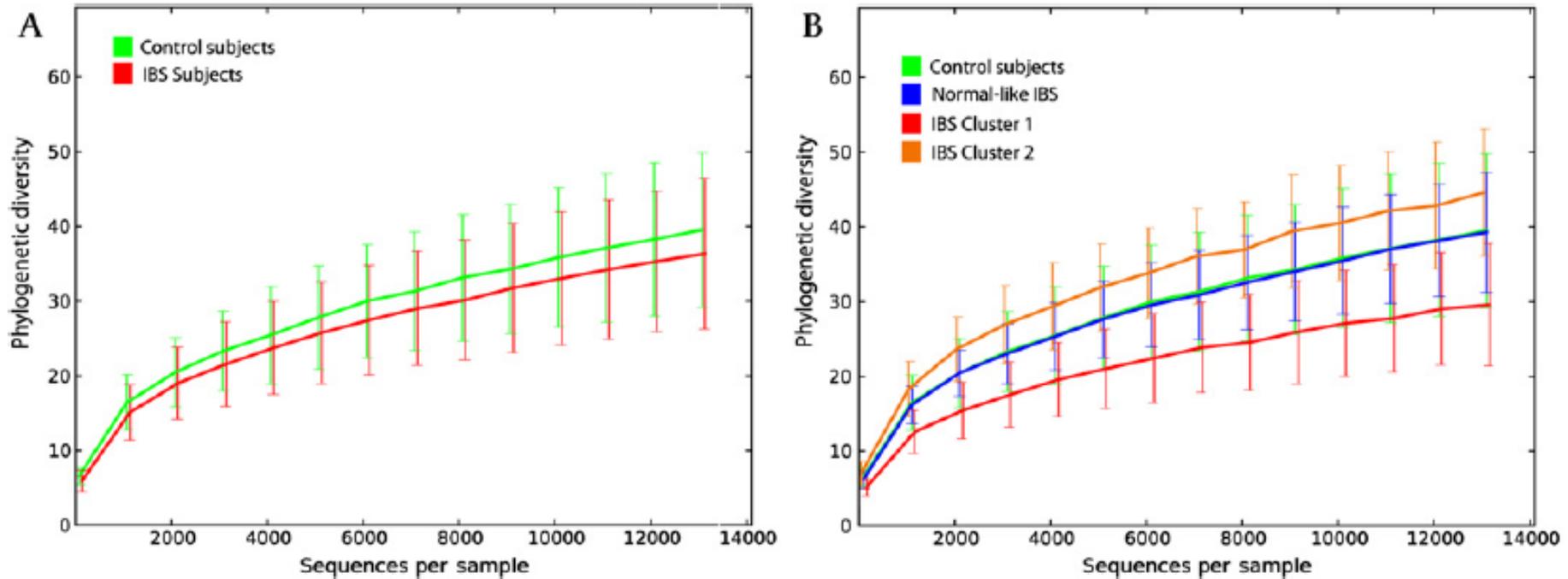


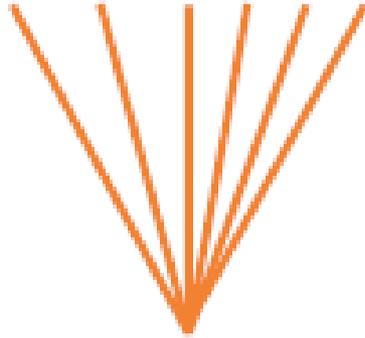
Figure 2 Faith's phylogenetic diversity curves. α Diversity curves showing total branch length with SDs for each group, for each number of sequences generated through rarefaction for (A) control and irritable bowel syndrome (IBS) samples and (B) control and IBS subgroups.

An irritable bowel syndrome subtype defined by species-specific alterations in faecal microbiota.

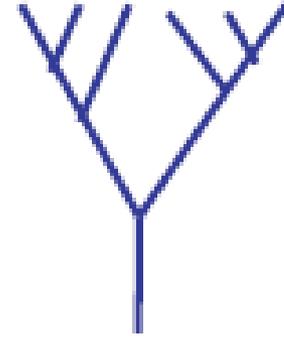
Gut 2012;61:997e1006

But perhaps
the slope will
vary.....

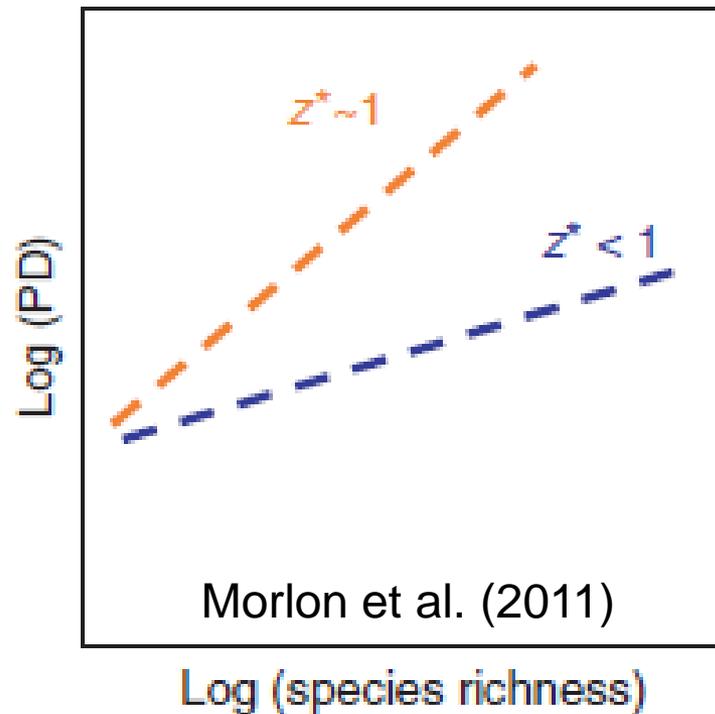
High
distinctiveness



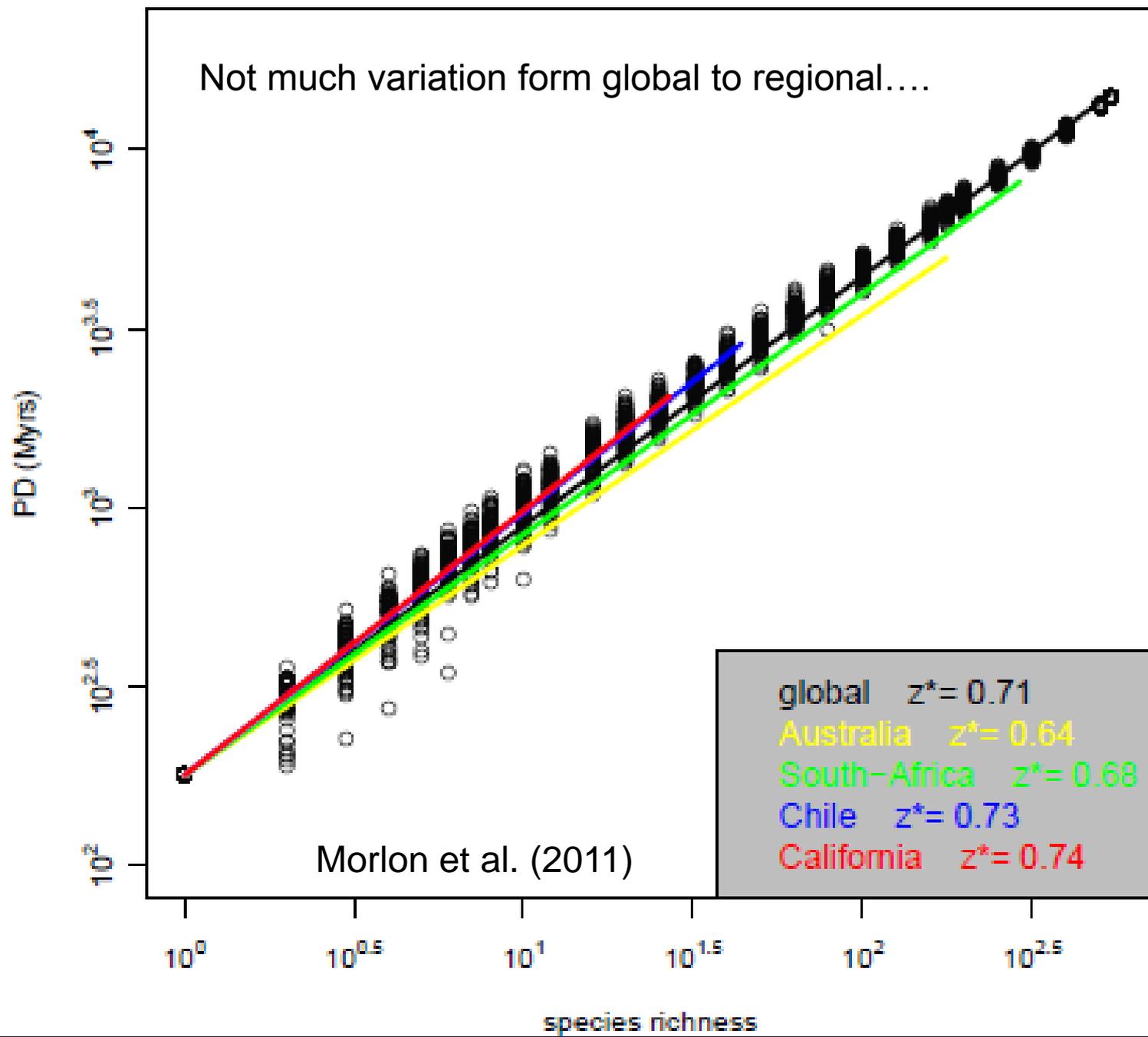
Low
distinctiveness



(a)



Not much variation from global to regional....



Morlon et al. (2011)

global $z^* = 0.71$
Australia $z^* = 0.64$
South-Africa $z^* = 0.68$
Chile $z^* = 0.73$
California $z^* = 0.74$

Interested in the departures from this PD-species relationship

- Residuals from modelled relationship have geographic/habitat pattern
- Clumped impacts on the tree
- Other PD calculations differ from corresponding species calculations
 - Complementarity
 - endemism

Interested in the departures from this PD-species relationship

- Residuals from modelled relationship have geographic/habitat pattern
- Clumped impacts on the tree
- Other PD calculations differ from corresponding species calculations
 - Complementarity
 - endemism

Mammal species richness map (top) looks a lot like mammal PD map (bottom)

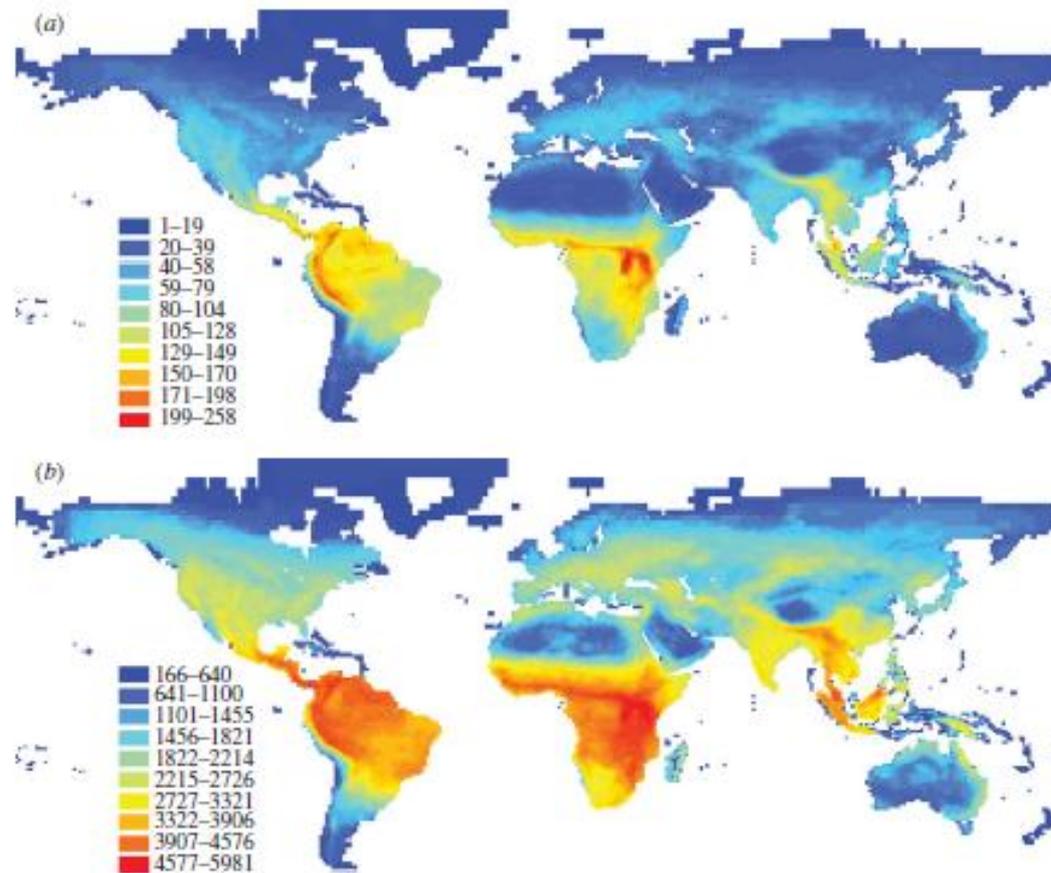
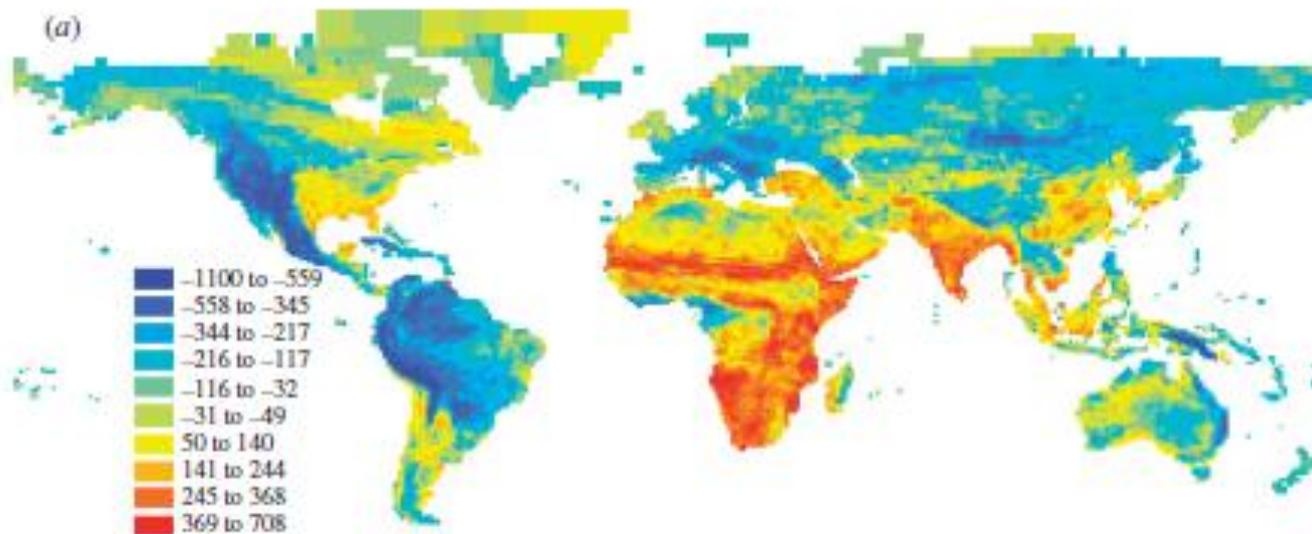


Figure 1. (a) Mammal species richness and (b) mammal phylogenetic diversity (millions of years). Data are depicted divided into 32 classes using natural breaks.

But the residuals from a model linking PD and species richness shows geographic patterns

2420 T. J. Davies & L. B. Buckley *Phylogenetic diversity of mammals*



Reductions in phylogenetic diversity in southern Europe, gains in north

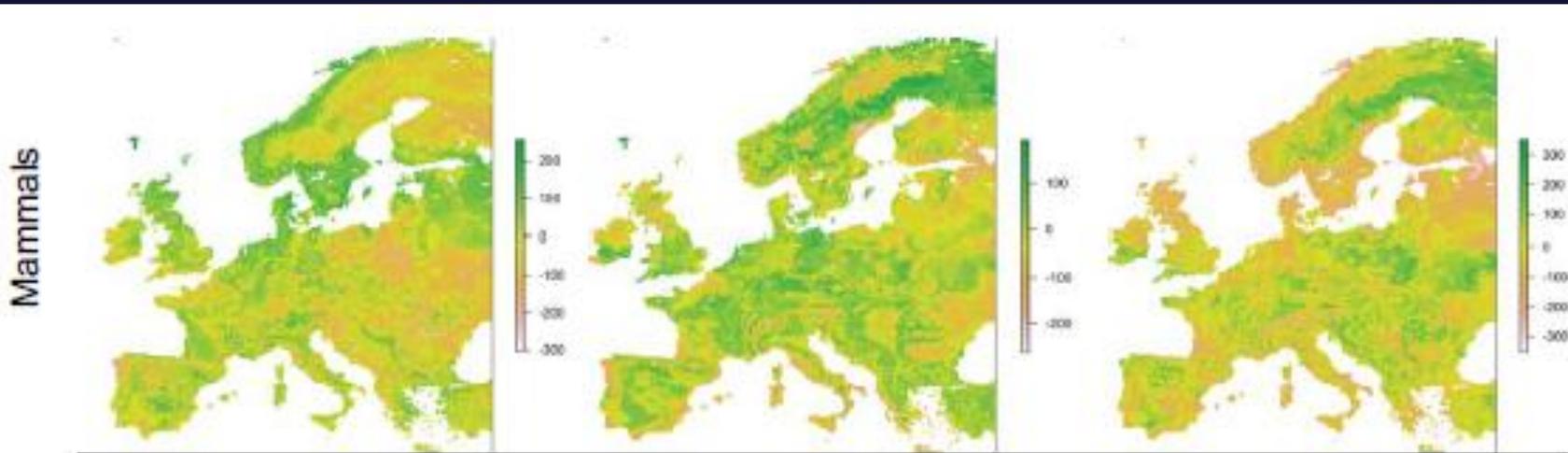


Figure S8: Predicted change in phylogenetic diversity under scenario A1 HadCM3 for 2080 after accounting for species richness. To make sure the change in the spatial phylogenetic diversity was not only due to the change in species richness, we mapped the residuals of a generalised additive model with phylogenetic diversity as the response variable and species richness as the explanatory variable under both current and future conditions. A) Map of residuals from the generalised additive regression model from current phylogenetic diversity and species richness. When positive, phylogenetic diversity is higher than expected given the mean species richness, and reciprocally. B) Map of residuals from the generalised additive regression model from predicted future phylogenetic diversity and future species richness; C) Difference between current residuals (B) and future residuals (A). When positive, change in phylogenetic diversity is higher than expected given change in species richness, and reciprocally. For plants and birds, the spatial variation in alpha PD is the average PD across the 100 trees.

Thuiller et al (2011) Nature 470, 531-534.

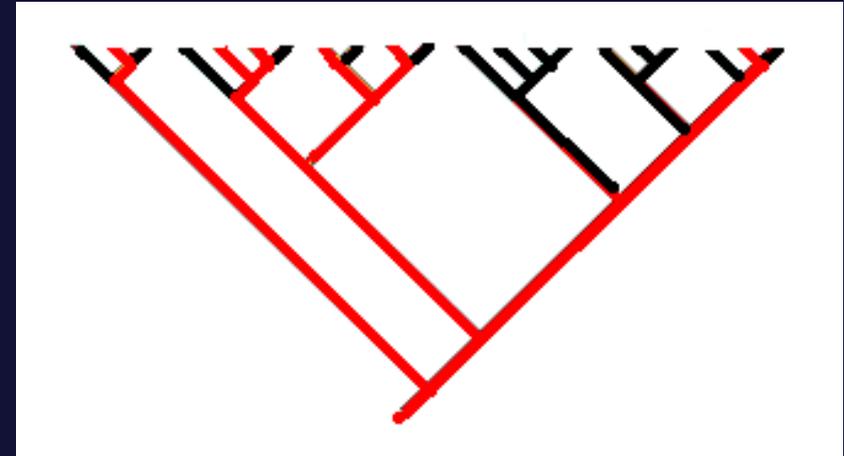
Interested in the departures from this PD-species relationship

- Residuals from modelled relationship have geographic/habitat pattern
- **Clumped impacts on the tree**
- Other PD calculations differ from corresponding species calculations
 - Complementarity
 - endemism

Will the impacts of climate change on PD be large or small?

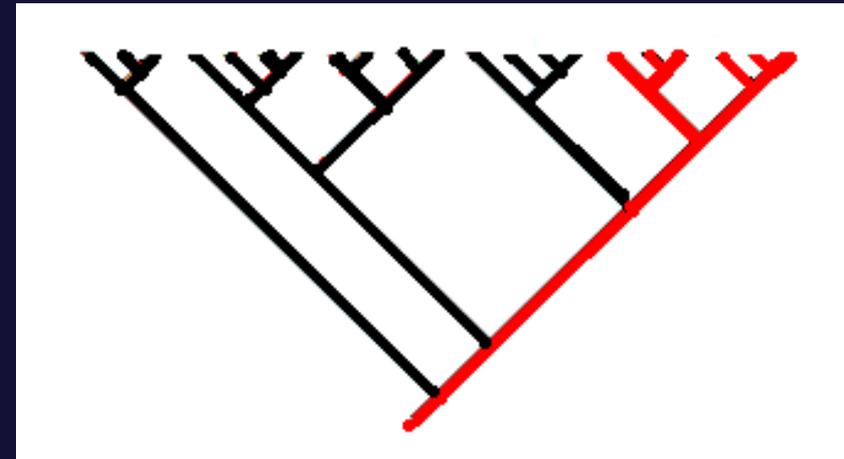
Yesson, C. and A. Culham. 2006.

- *small* loss of PD or evolutionary potential for given species loss



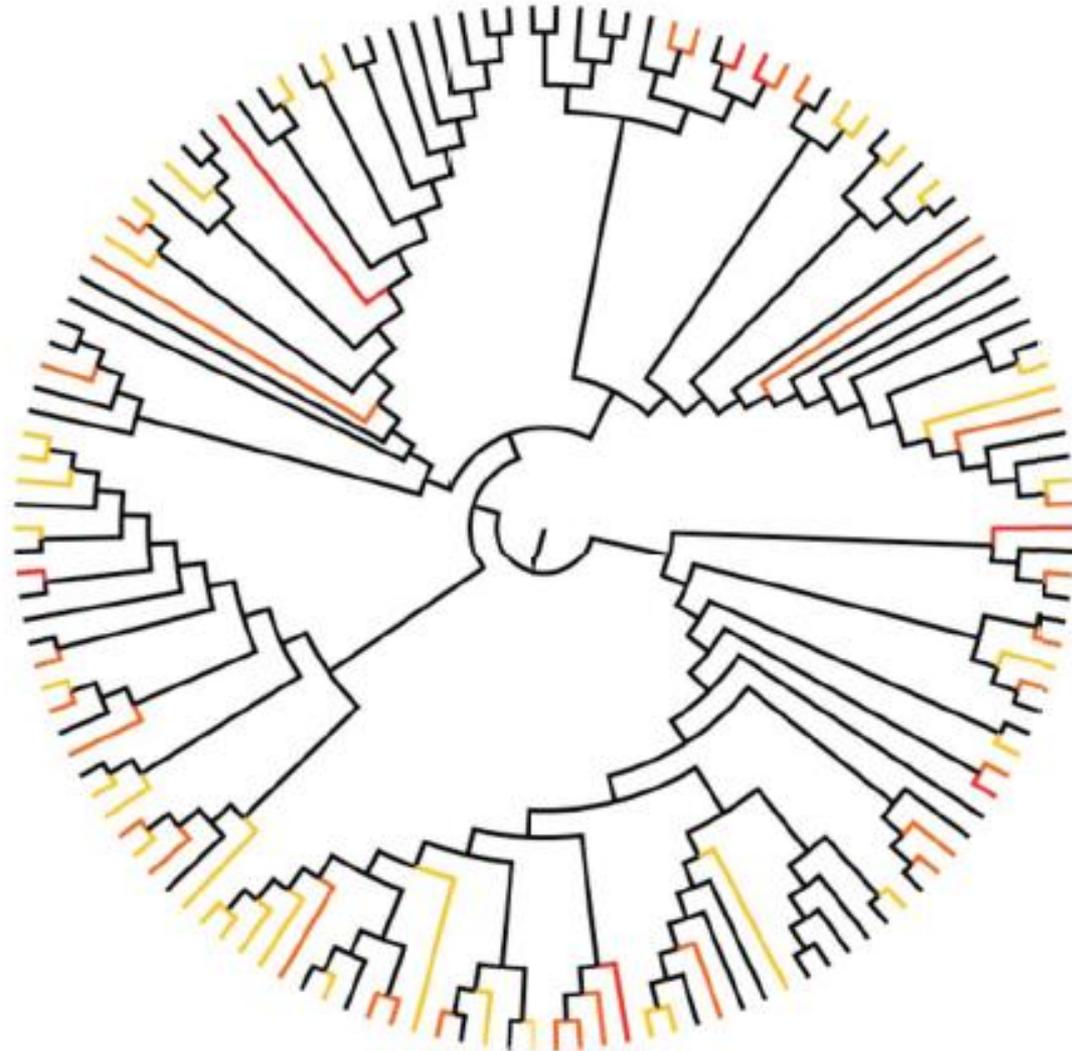
red = surviving evolutionary potential

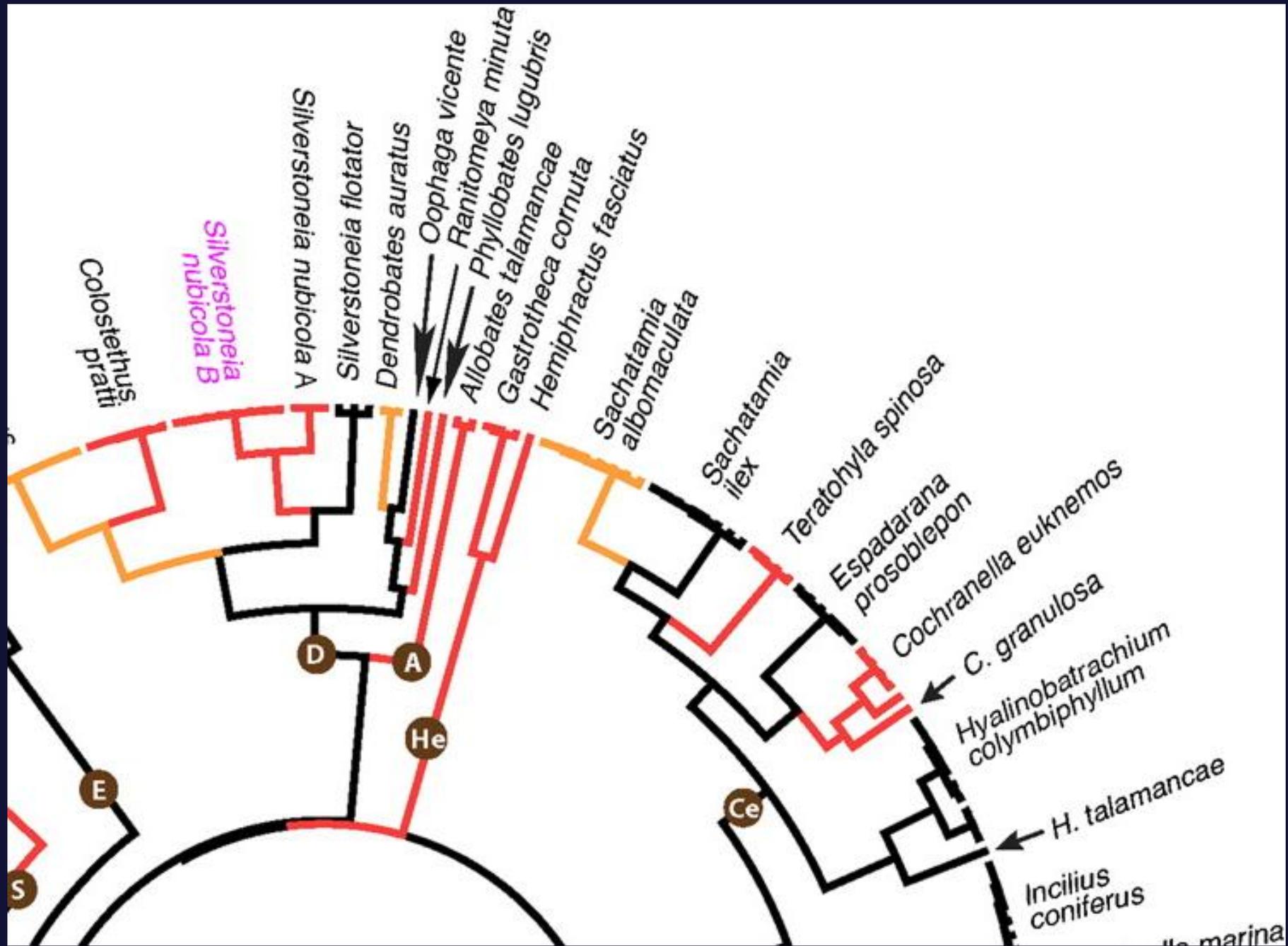
- *large* loss of PD or evolutionary potential



Percent loss of habitat, colour coded on tree Thuiller et al 2011

Figure S5 C – Mammals

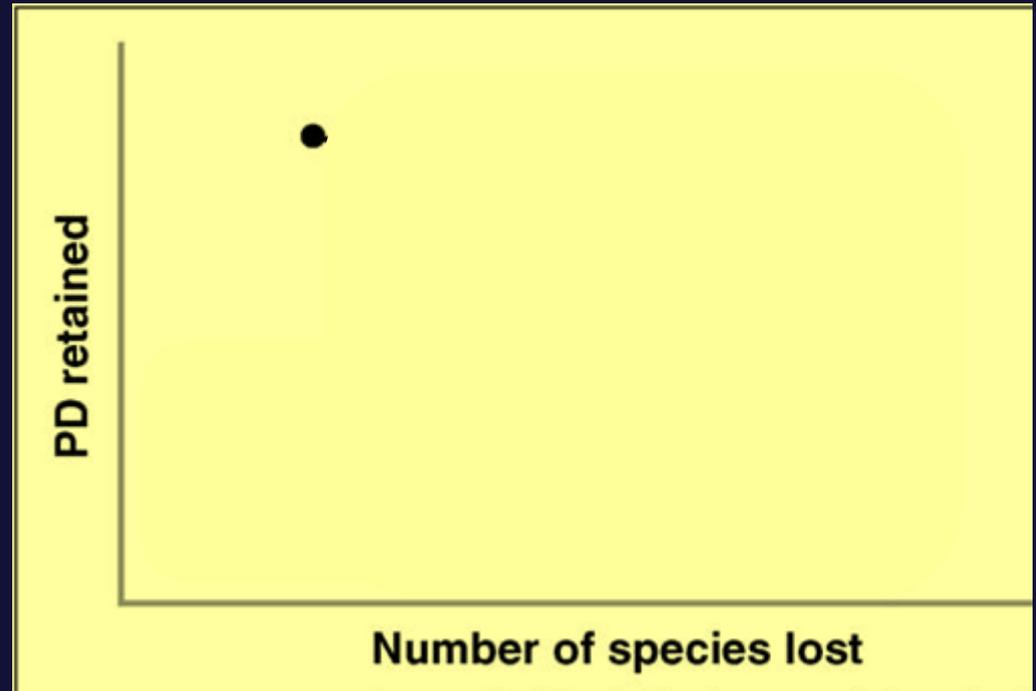
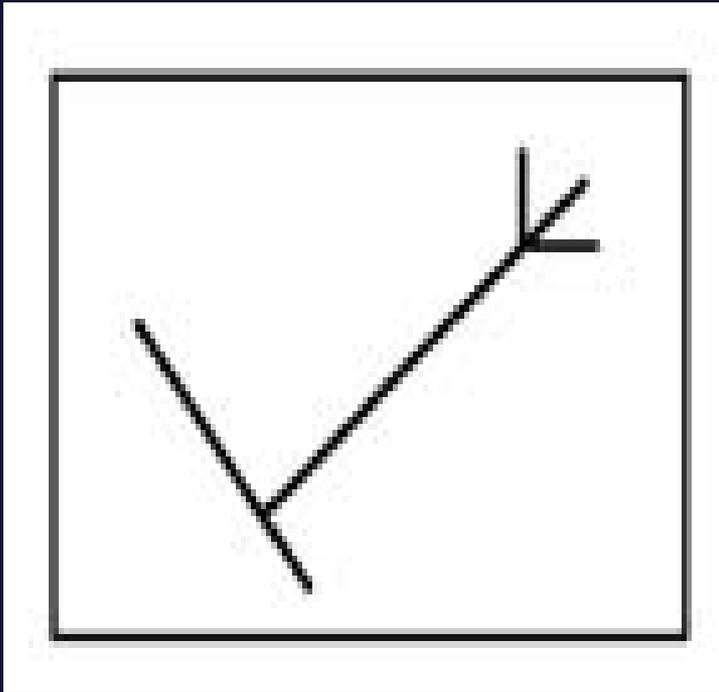




H. marina

Phylogenetic diversity and tipping points

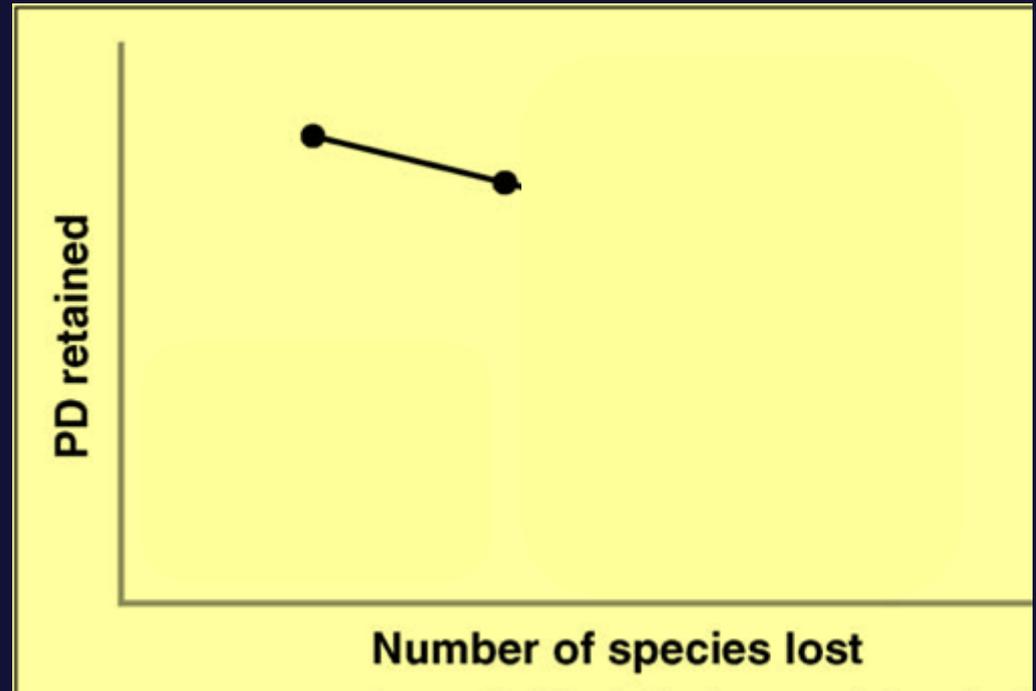
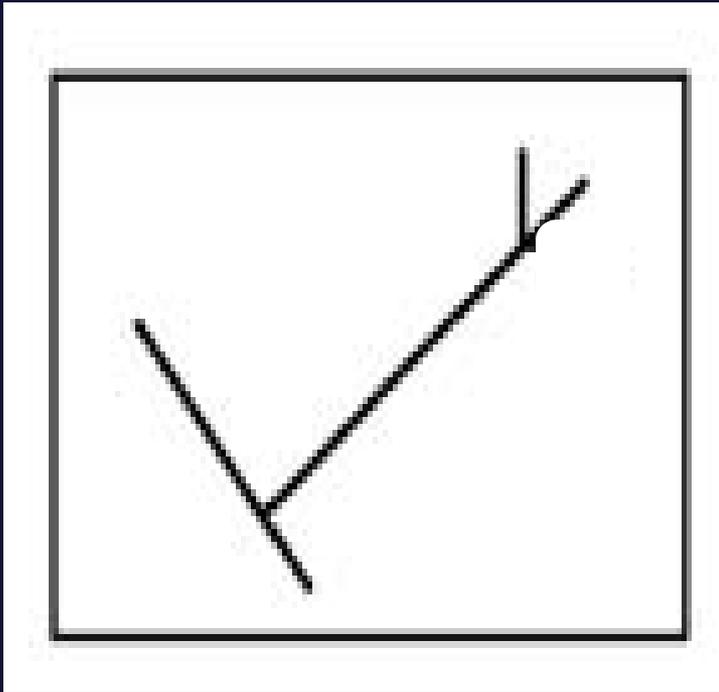
PD loss (vertical axis) as species are lost (horizontal axis)
Loss of one species, and loss of a second species imply small PD losses, but loss of the third species is a tipping point — the deeper ancestral branch and corresponding PD is now lost.



Faith et al (2010) Ecosystem services: an evolutionary perspective on the links between biodiversity and human well-being. *COSUST*

Phylogenetic diversity and tipping points

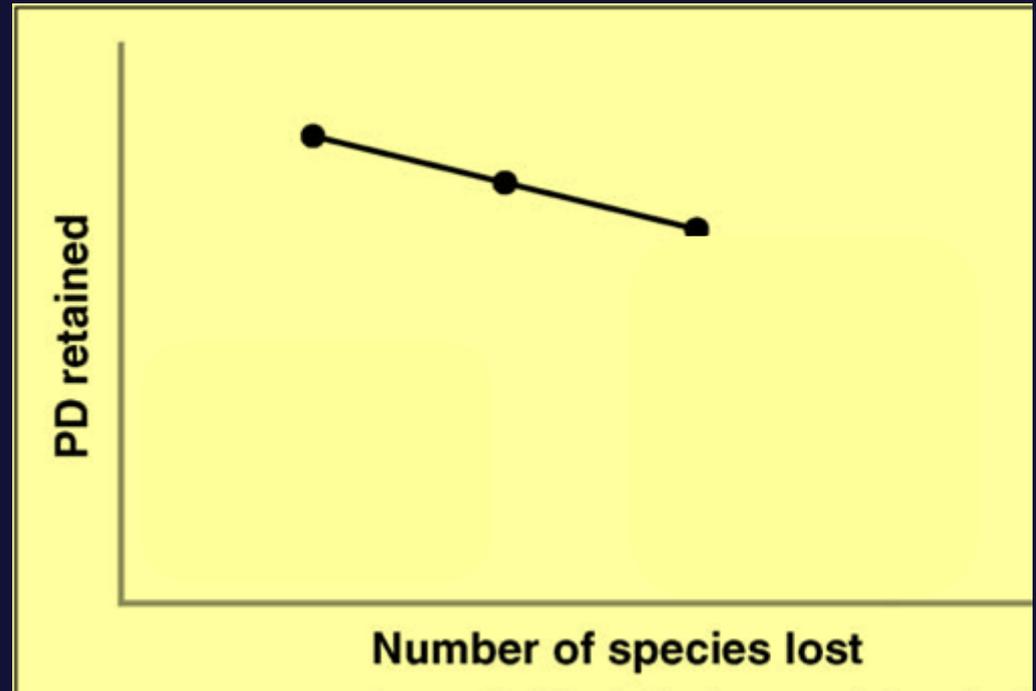
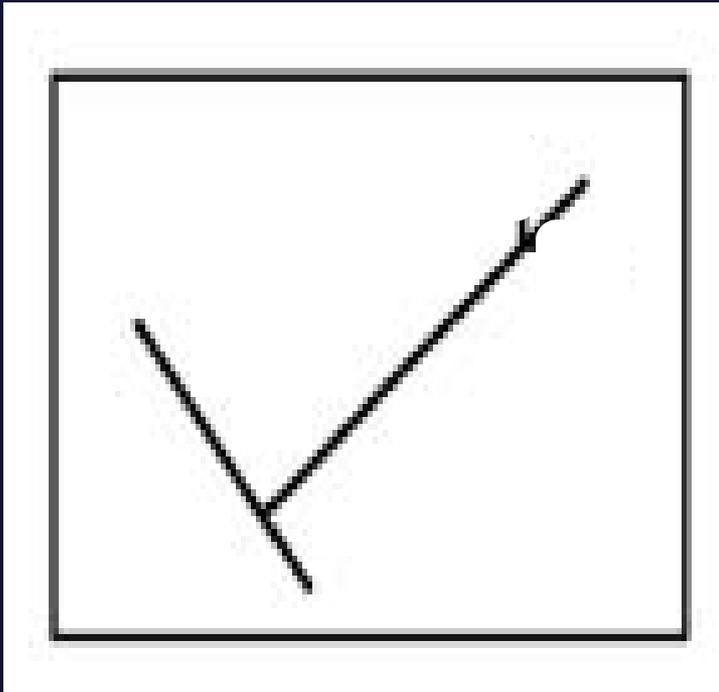
PD loss (vertical axis) as species are lost (horizontal axis)
Loss of one species, and loss of a second species imply small PD losses, but loss of the third species is a tipping point — the deeper ancestral branch and corresponding PD is now lost.



Faith et al (2010) Ecosystem services: an evolutionary perspective on the links between biodiversity and human well-being. *COSUST*

Phylogenetic diversity and tipping points

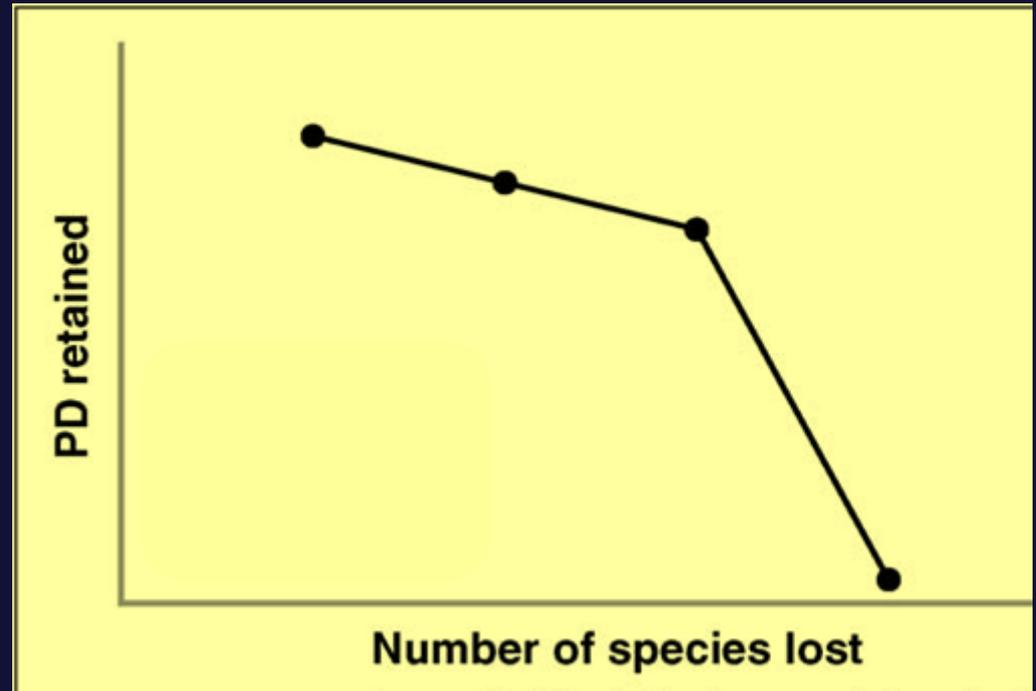
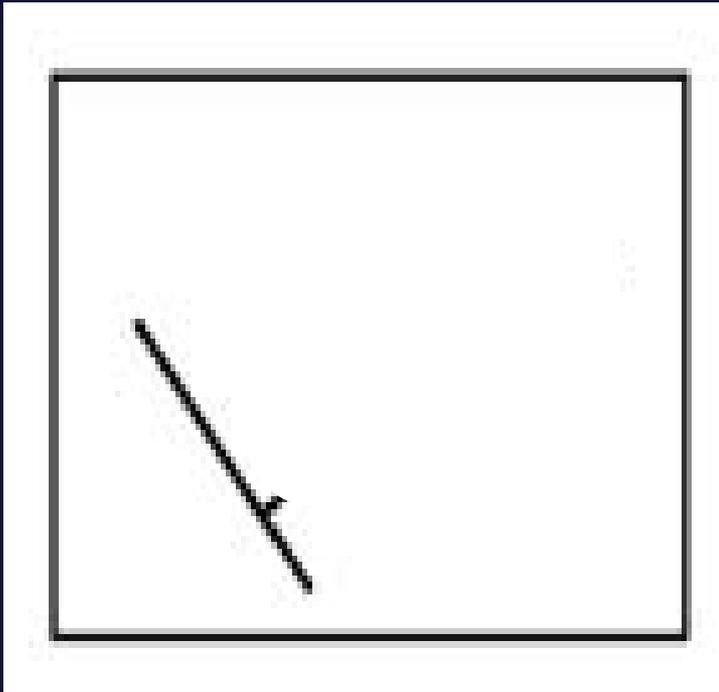
PD loss (vertical axis) as species are lost (horizontal axis)
Loss of one species, and loss of a second species imply small PD losses, but loss of the third species is a tipping point — the deeper ancestral branch and corresponding PD is now lost.



Faith et al (2010) Ecosystem services: an evolutionary perspective on the links between biodiversity and human well-being. *COSUST*

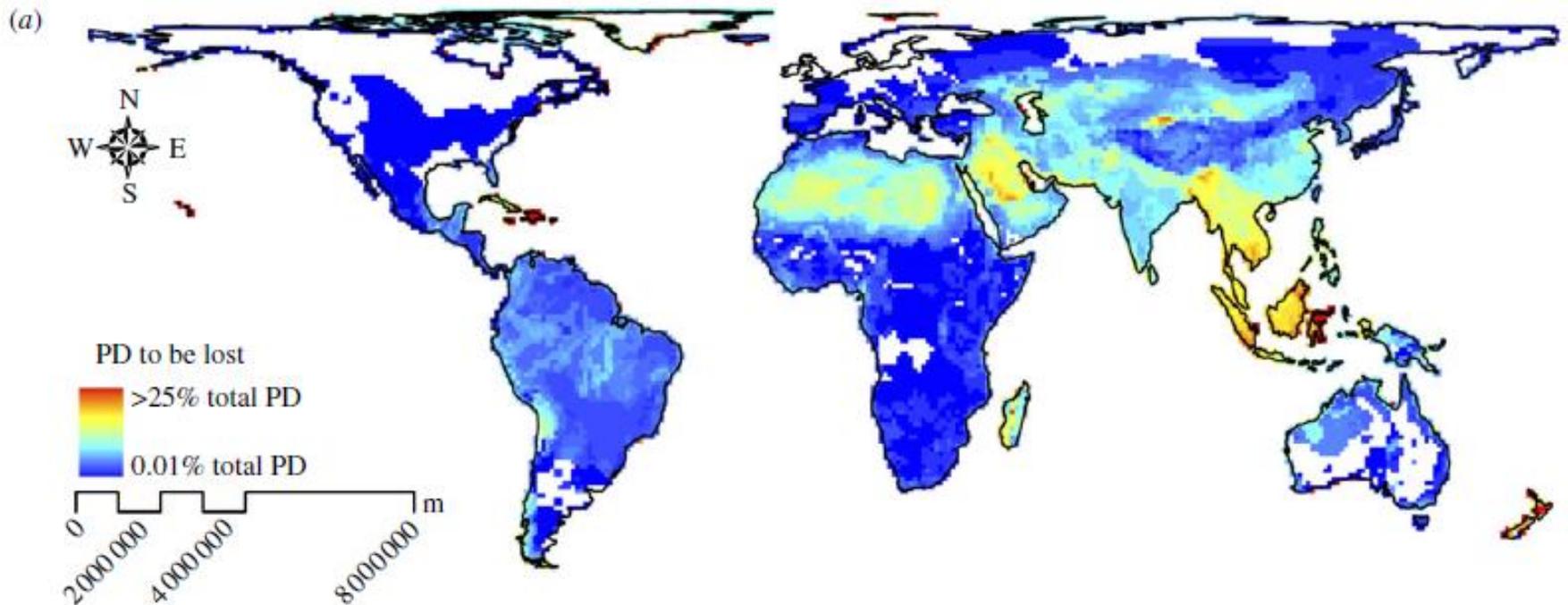
Phylogenetic diversity and tipping points

PD loss (vertical axis) as species are lost (horizontal axis)
Loss of one species, and loss of a second species imply small PD losses, but loss of the third species is a tipping point — the deeper ancestral branch and corresponding PD is now lost.



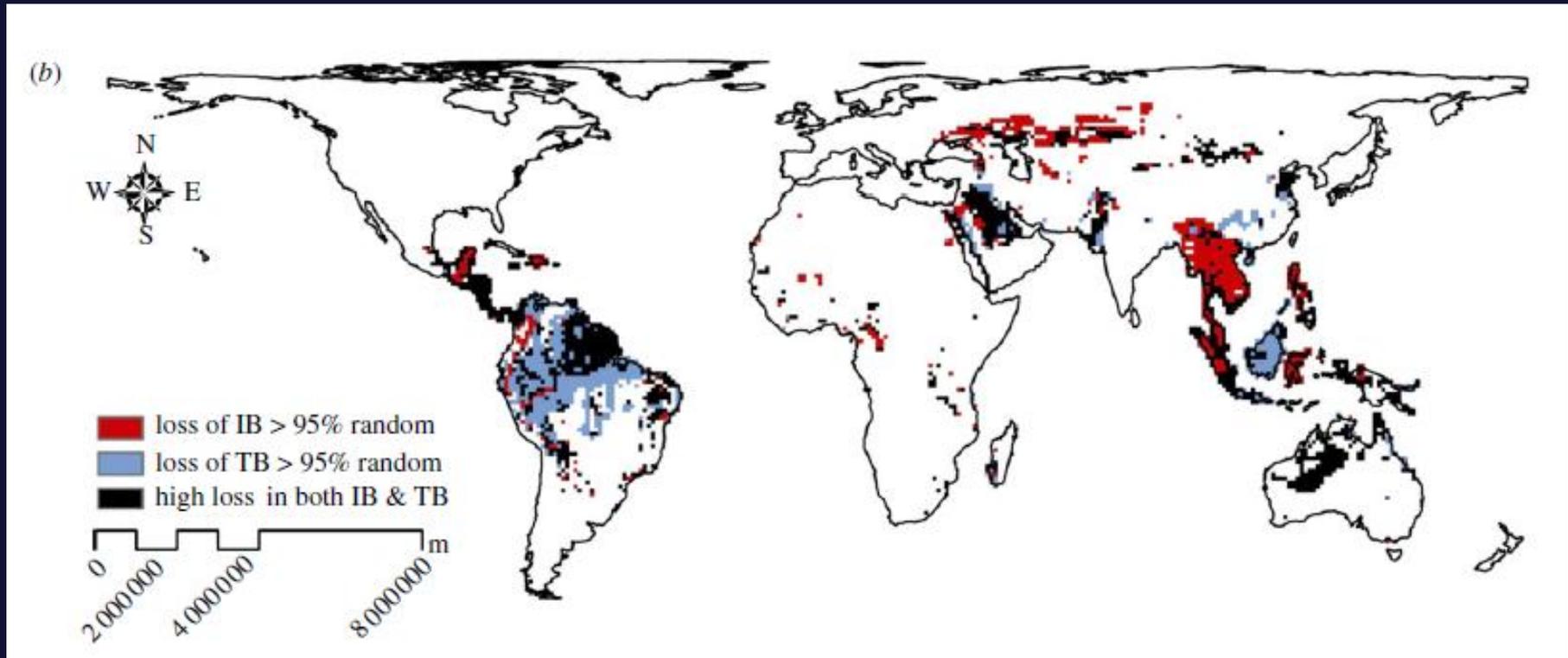
Faith et al (2010) Ecosystem services: an evolutionary perspective on the links between biodiversity and human well-being. *COSUST*

2 S. Huang et al. *Global extinctions, regional impacts*



Predicted phylogenetic diversity (PD) to be lost with extinction of currently threatened species within each 100 by 100 km grid cell. The colours represent the percentage of PD to be lost, from low (blue) to high (red).

Comparison with random losses reveals clumped impacts and possible tipping points. Red areas have high loss of deeper branches on phylogeny of mammals.

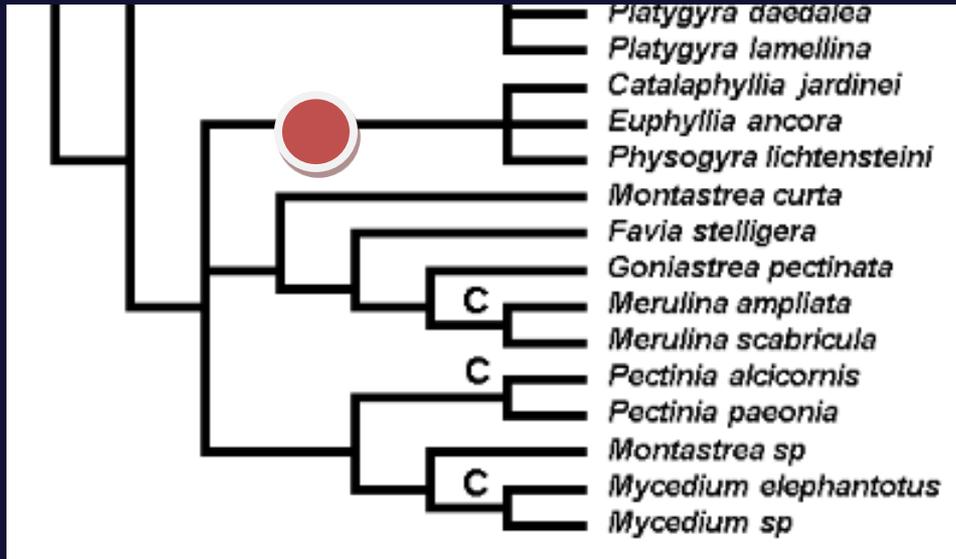


“At regional scales, losses differ dramatically: several biodiversity hotspots in southern Asia and Amazonia will lose an unexpectedly large proportion of PD.”

Huang et al (2011) Biol. Lett.

Loss of the world's corals

- “the proportion of corals (57.8%) exceeds that of all terrestrial animal groups assessed to date..”
- Carpenter et al (2008) *Science*



sometimes entire clades fall into IUCN threatened classes

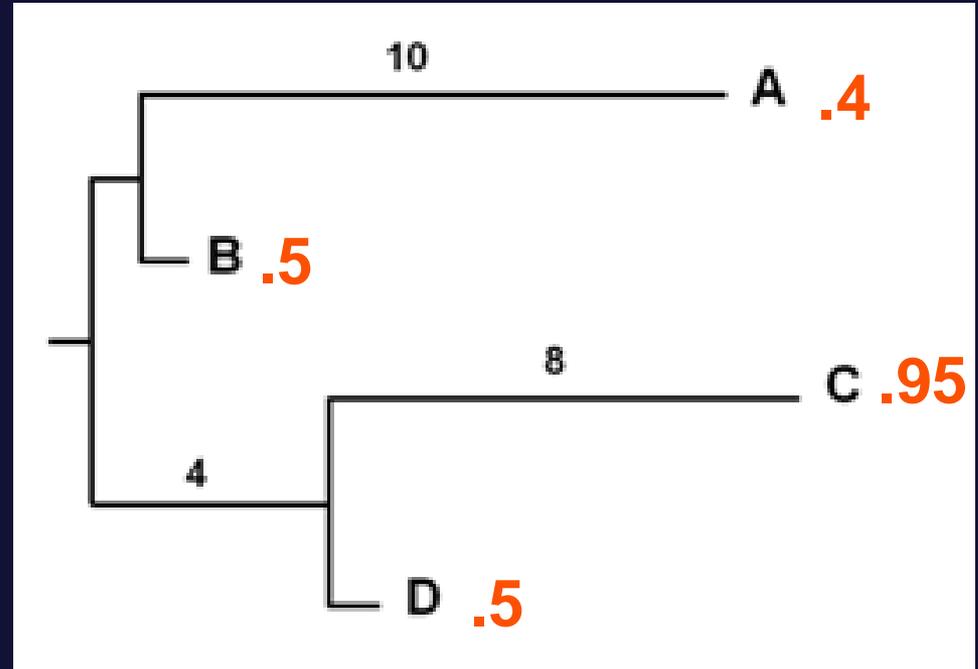
phylogenetic risk analysis

- Faith DP and ZT Richards (to appear) Implications of climate change for the tree of life. *Biology*

PD and probabilities of extinction

Probabilistic PD

- Red numbers are estimated probabilities of extinction



Can estimate “expected phylogenetic diversity”
or do “phylogenetic risk analysis”

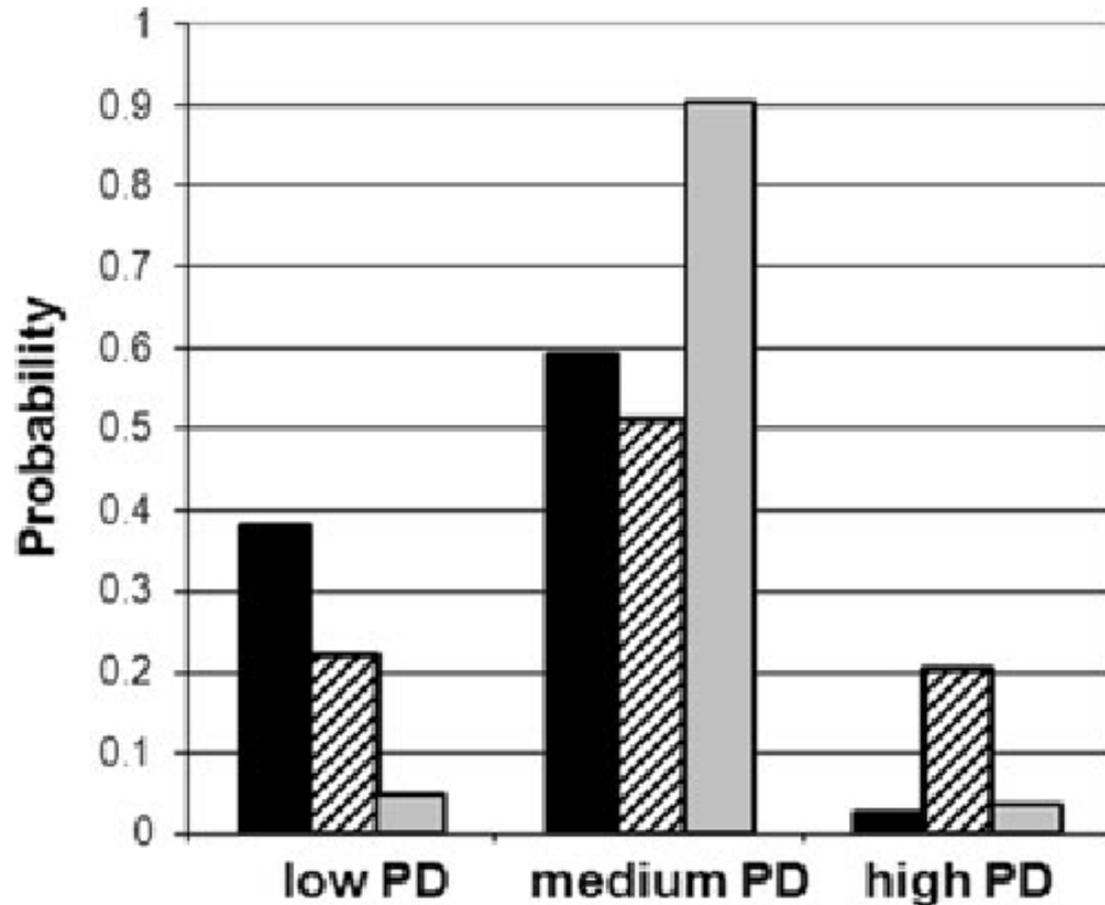
Faith DP (2008) Threatened species and the preservation of phylogenetic diversity (PD): assessments based on extinction probabilities and risk analysis. *Conservation Biology*

Phylogenetic risk analysis

Black =
current

Striped =
protect
species to
max
expected PD

Gray =
select
species to
avoid worst
case losses



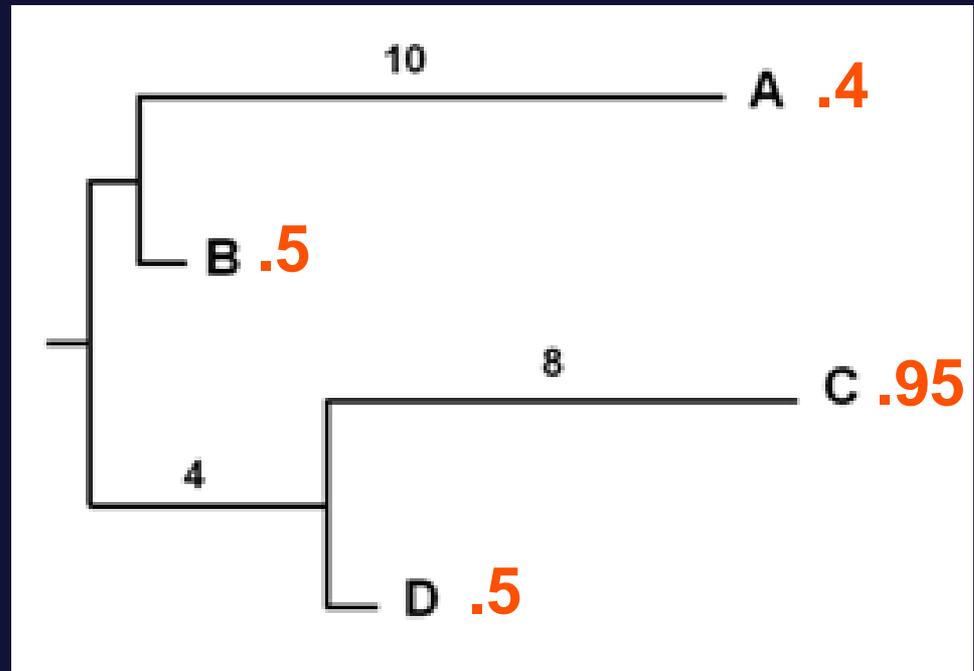
Faith DP (2008) *Conservation Biology*

Faith DP (2009) Phylogenetic triage, efficiency, and risk aversion.
Trends in Ecology and Evolution

PD and probabilities of extinction

Probabilistic PD

Red numbers
are estimated
probabilities of
extinction



Can estimate “expected phylogenetic diversity”

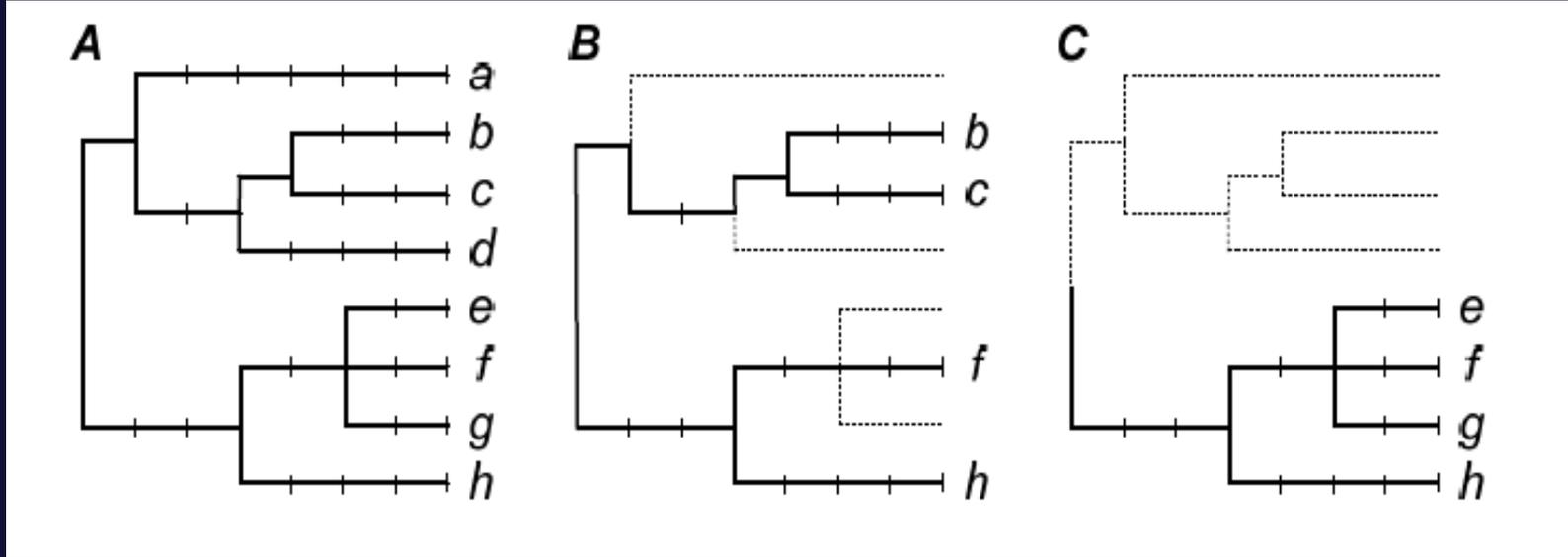
PD50 indices: for any species – what is the expected PD loss under extinction, assuming all other species have 50-50 chance of persistence? (see FISHBASE)

Interested in the departures from this PD-species relationship

- Residuals from modelled relationship have geographic/habitat pattern
- Clumped impacts on the tree
- Other PD calculations differ from corresponding species calculations
 - Complementarity
 - endemism

PD – phylogenetic diversity

Faith 1992



Phylogenetic ecology - take any conventional species-level index and re-express as a PD-based measure

Richness = total PD

Expected diversity = expected PD

PD-Complementarity (gains & losses)

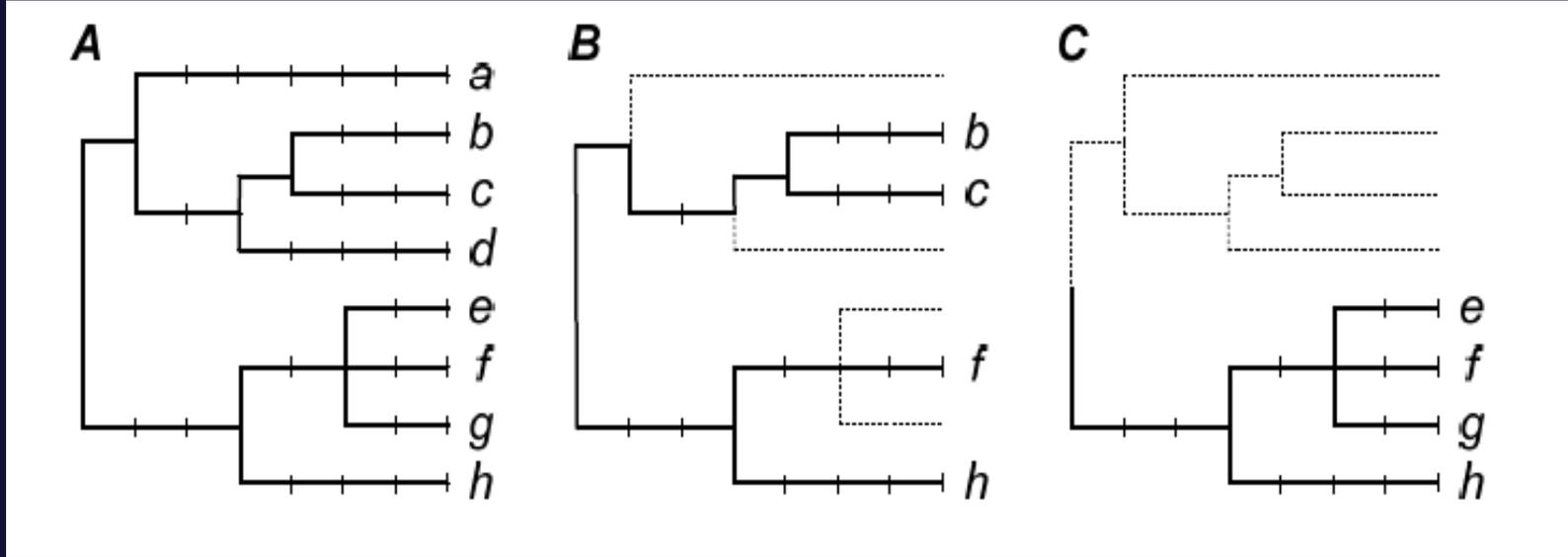
PD-Endemism (e.g. Faith et al 2004; Faith 1994)

PD-Dissimilarity between communities

PD analogues of Shannon-Weiner index and Simpson's index

PD – phylogenetic diversity

Faith 1992



Phylogenetic ecology - take any conventional species-level index and re-express as a PD-based measure

Richness = total PD

Expected diversity = expected PD

PD-Complementarity (gains & losses)

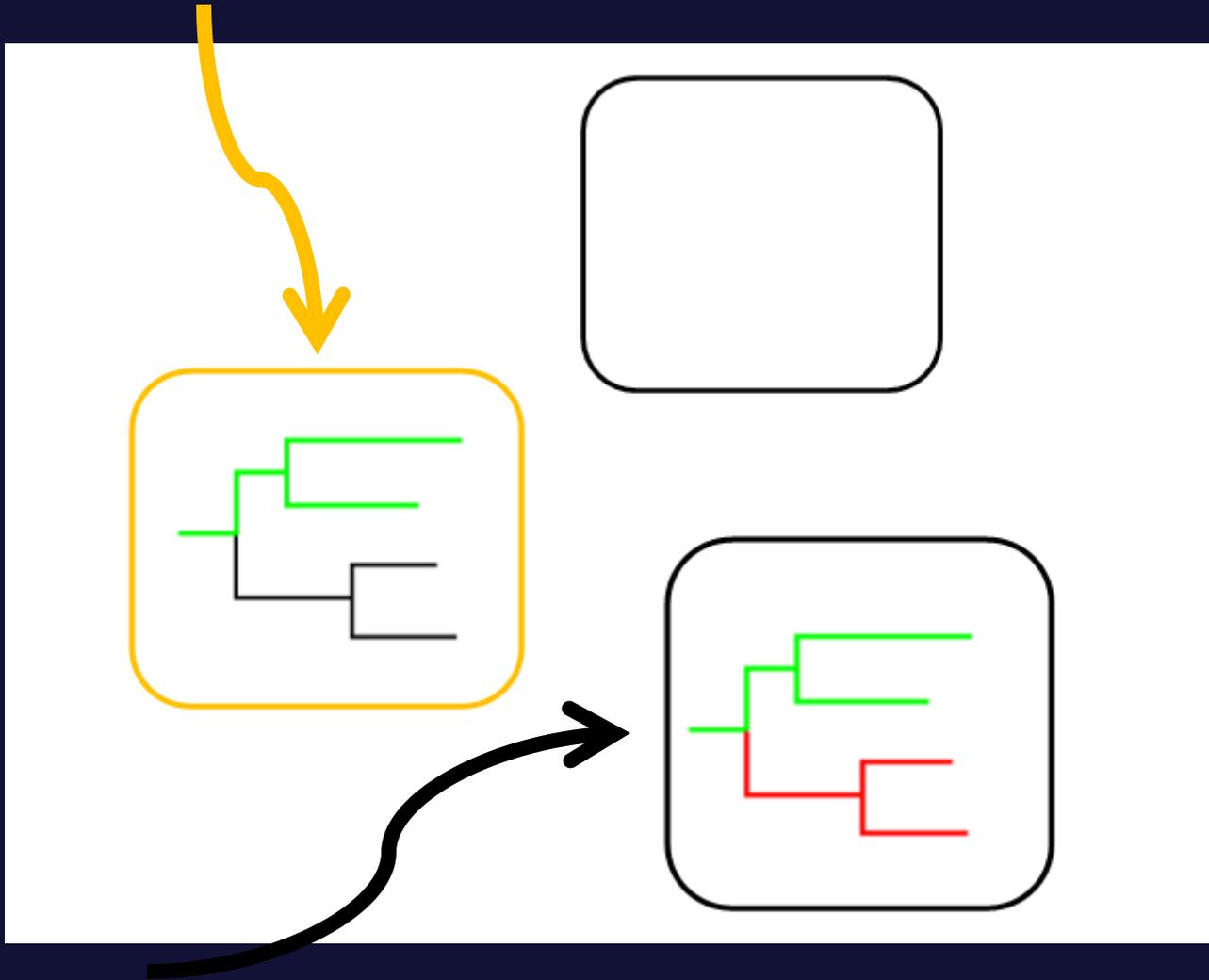
PD-Endemism (e.g. Faith et al 2004; Faith 1994)

PD-Dissimilarity between communities

PD analogues of Shannon-Weiner index and Simpson's index

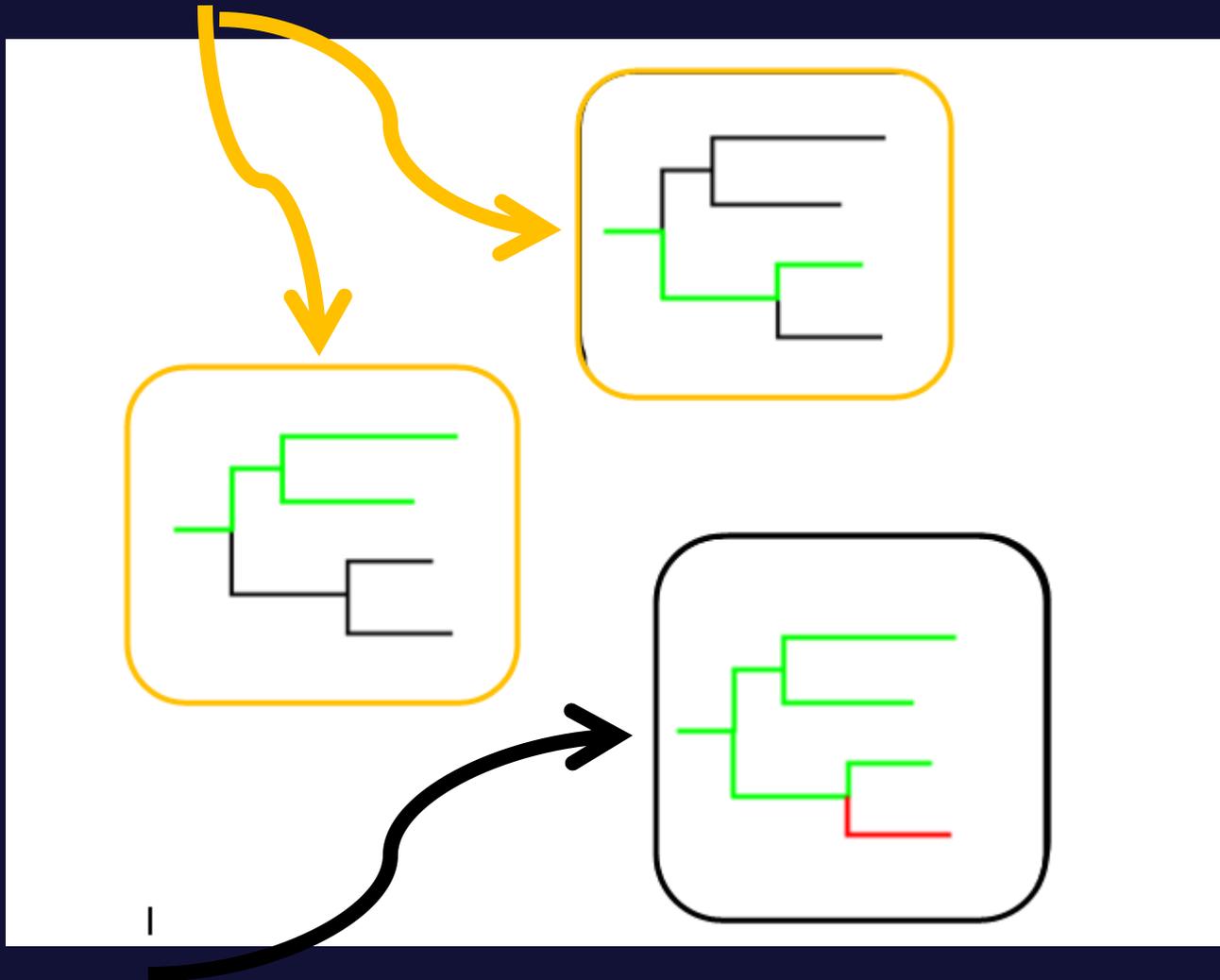
**Greater
differences
from species
calculations...**

Orange protected area, with protected PD in green



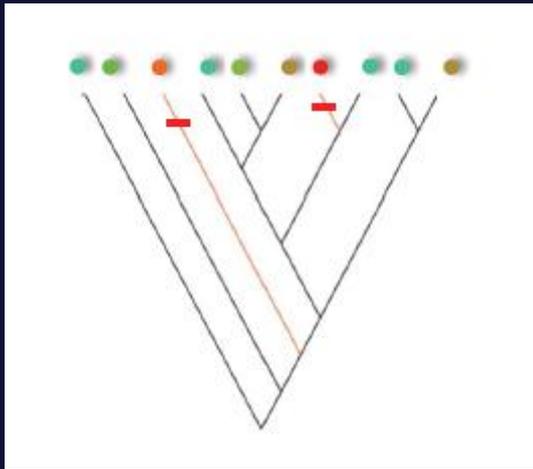
For this region, PD complementarity in red

Orange protected area, with protected PD in green

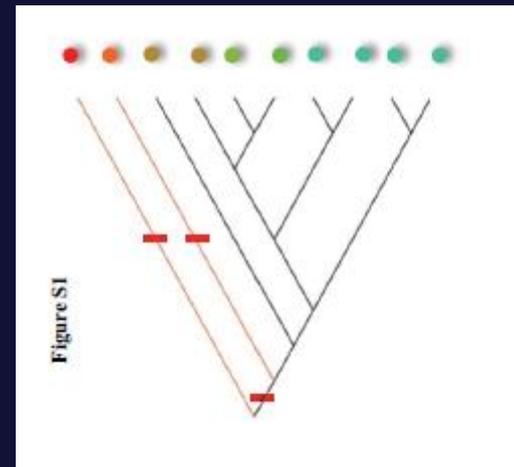
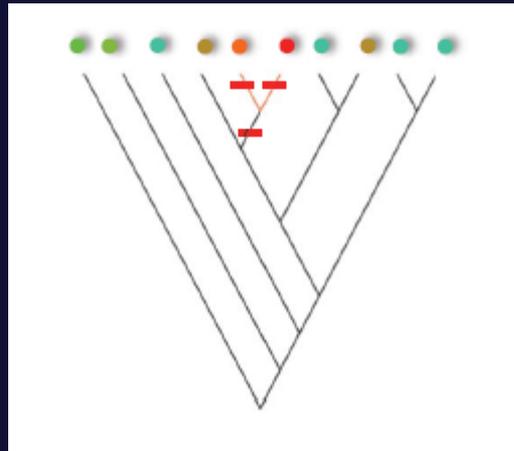


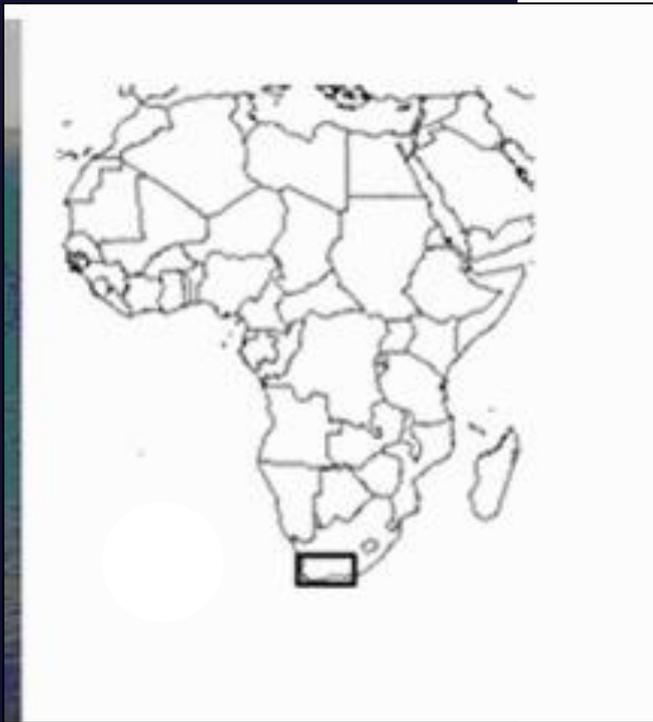
For this region, PD complementarity in red

Loss of two species can be large or small.....

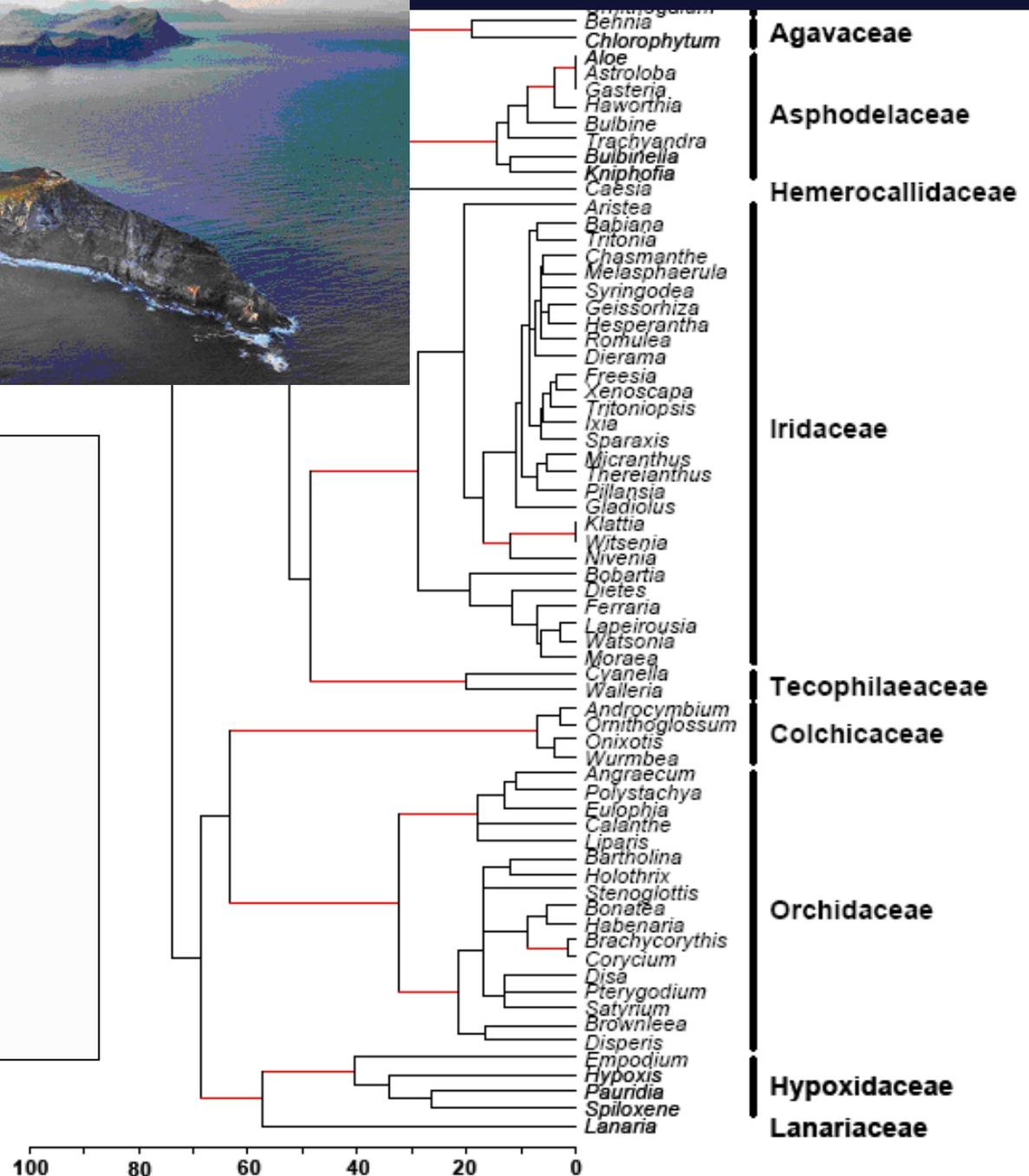


Thuiller et al 2011

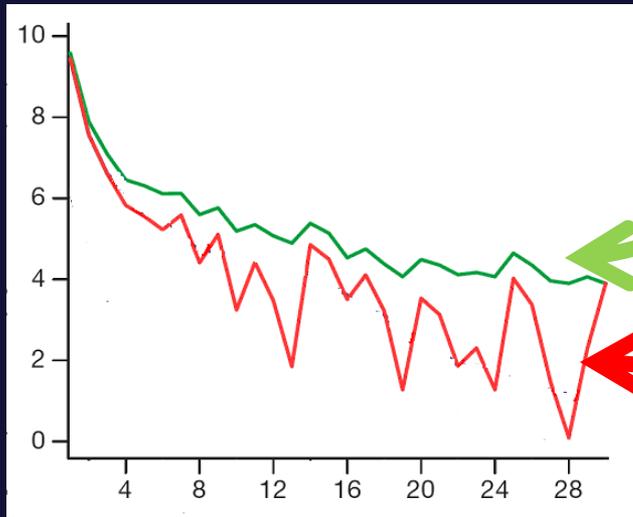




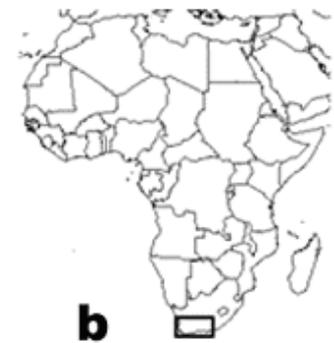
Forest et al. (2007)
Nature



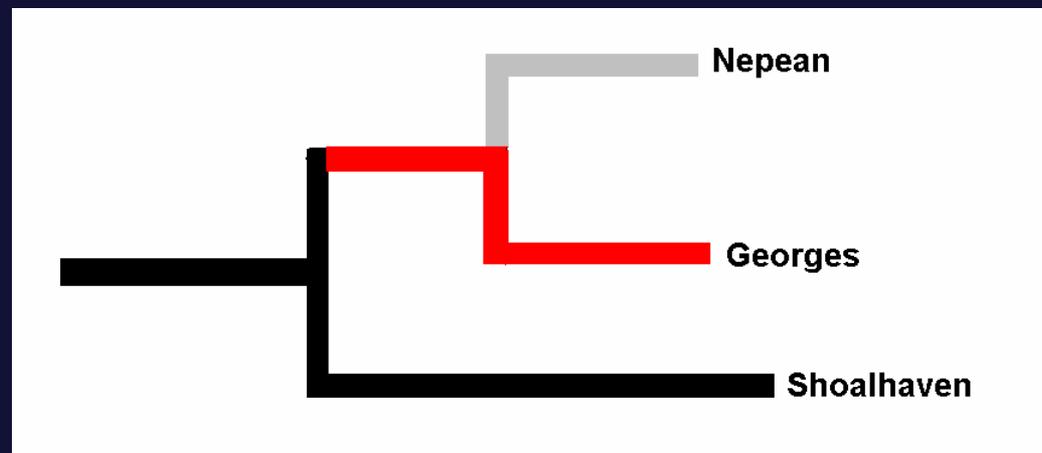
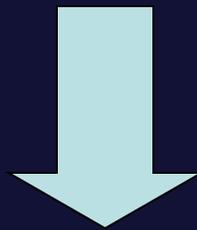
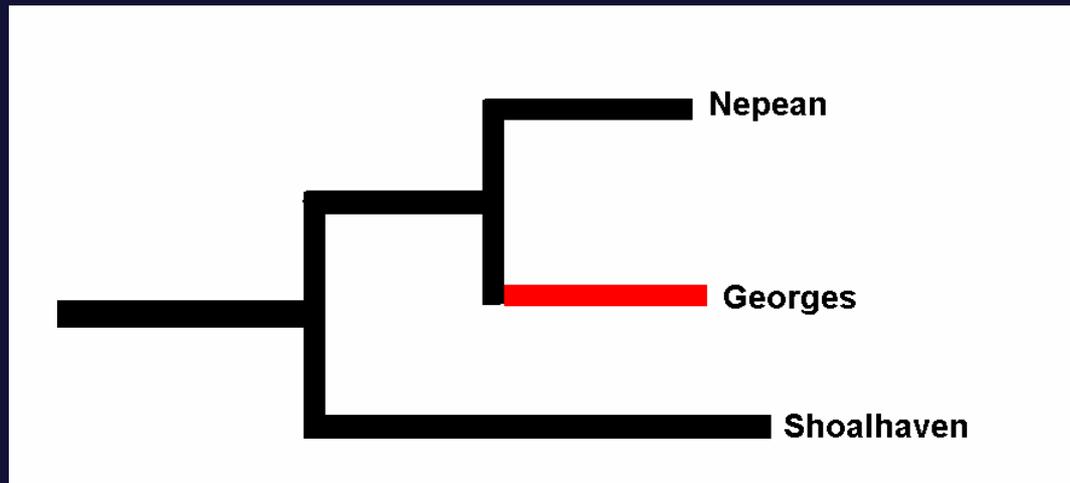
PD and the Cape hotspot: species counting highlights the western portion but PD highlights the eastern portion



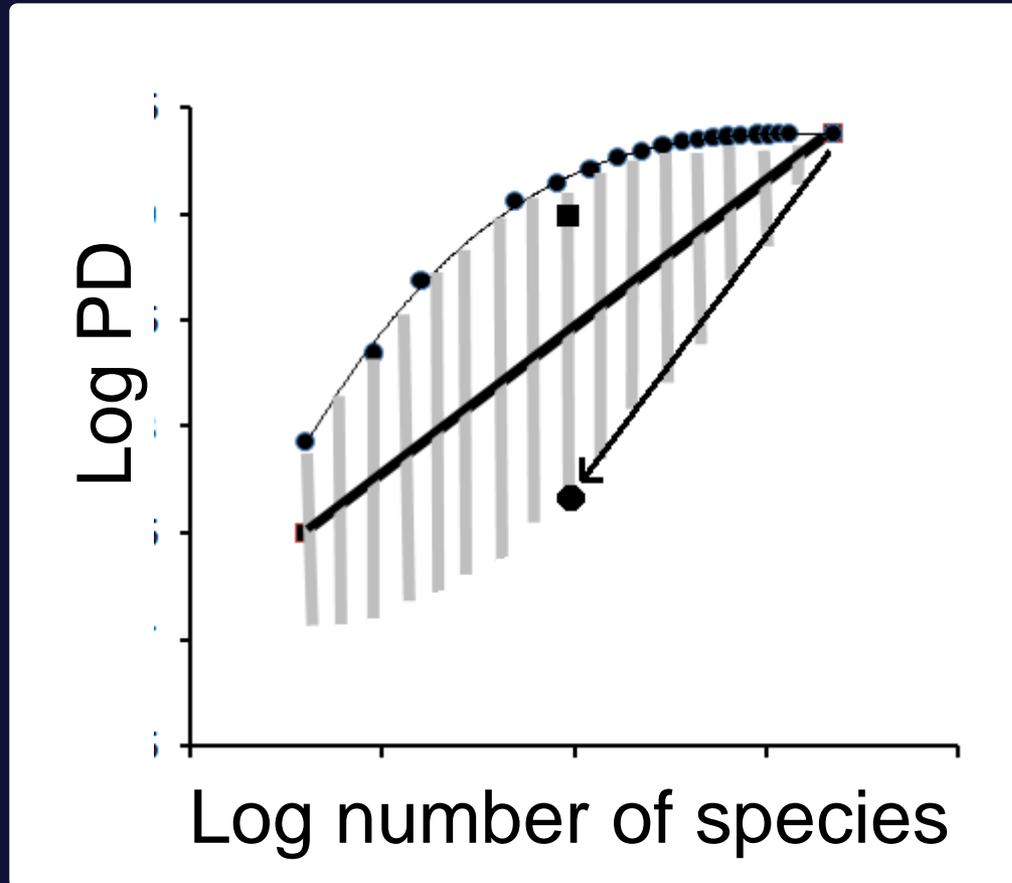
the PD that you could gain does not match the PD that you do gain



Forest et al
Nature 2007



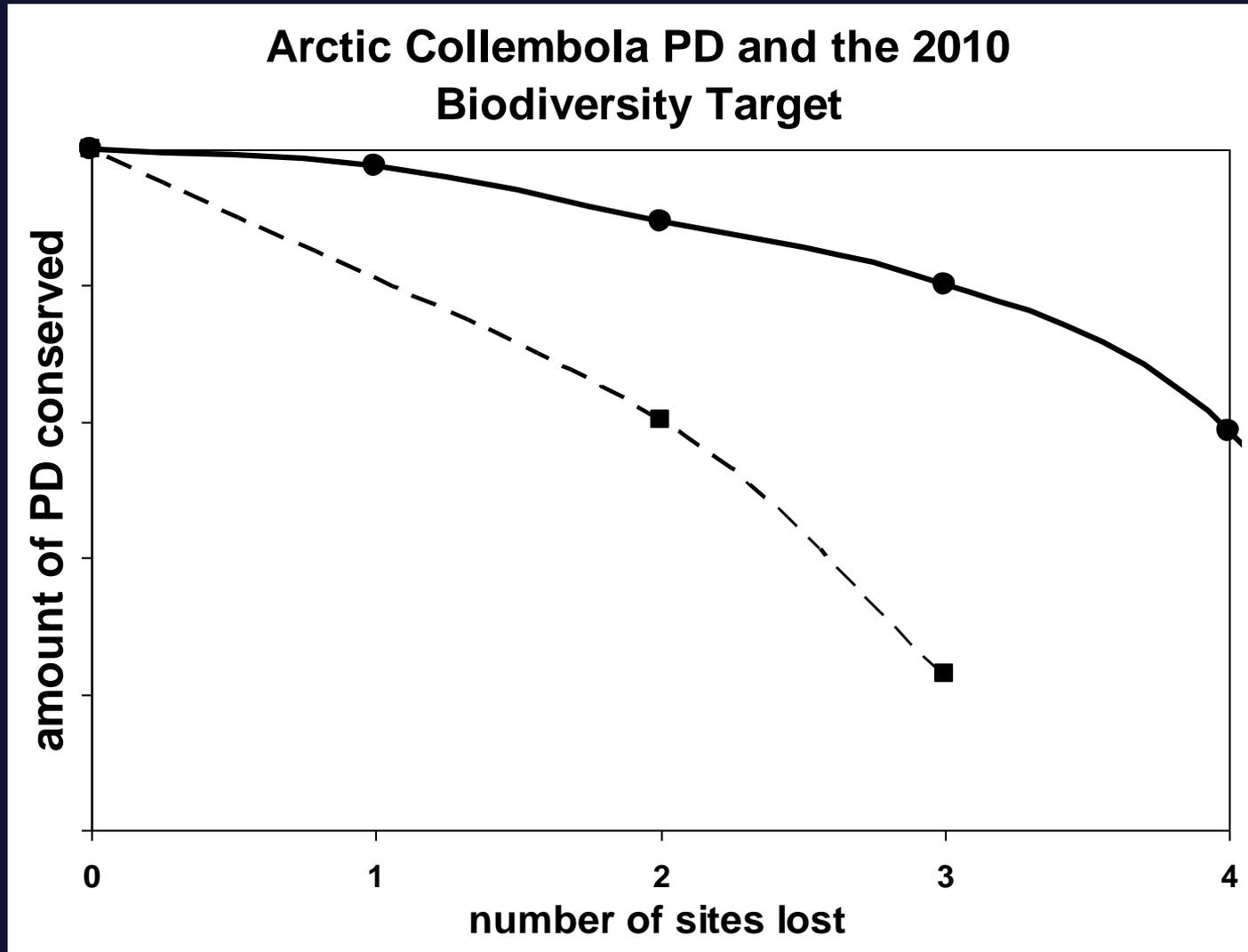
see also Faith, D. P. 2008. Phylogenetic diversity and conservation. In (eds: SP Carroll and C Fox) *Conservation Biology: Evolution in Action*. Oxford University Press.



We have the idealised PD – species curve (linear in log-log space)

But in reality optimised losses yield a curve higher up and perverse losses yield a curve further down

Prospects: a toolbox for application to “phylogenies” from large scale DNA barcoding programs



Faith, DP (2008) Phylogenetic diversity and conservation. In (eds: SP Carroll and C Fox) *Conservation Biology: Evolution in Action*. Oxford University Press.

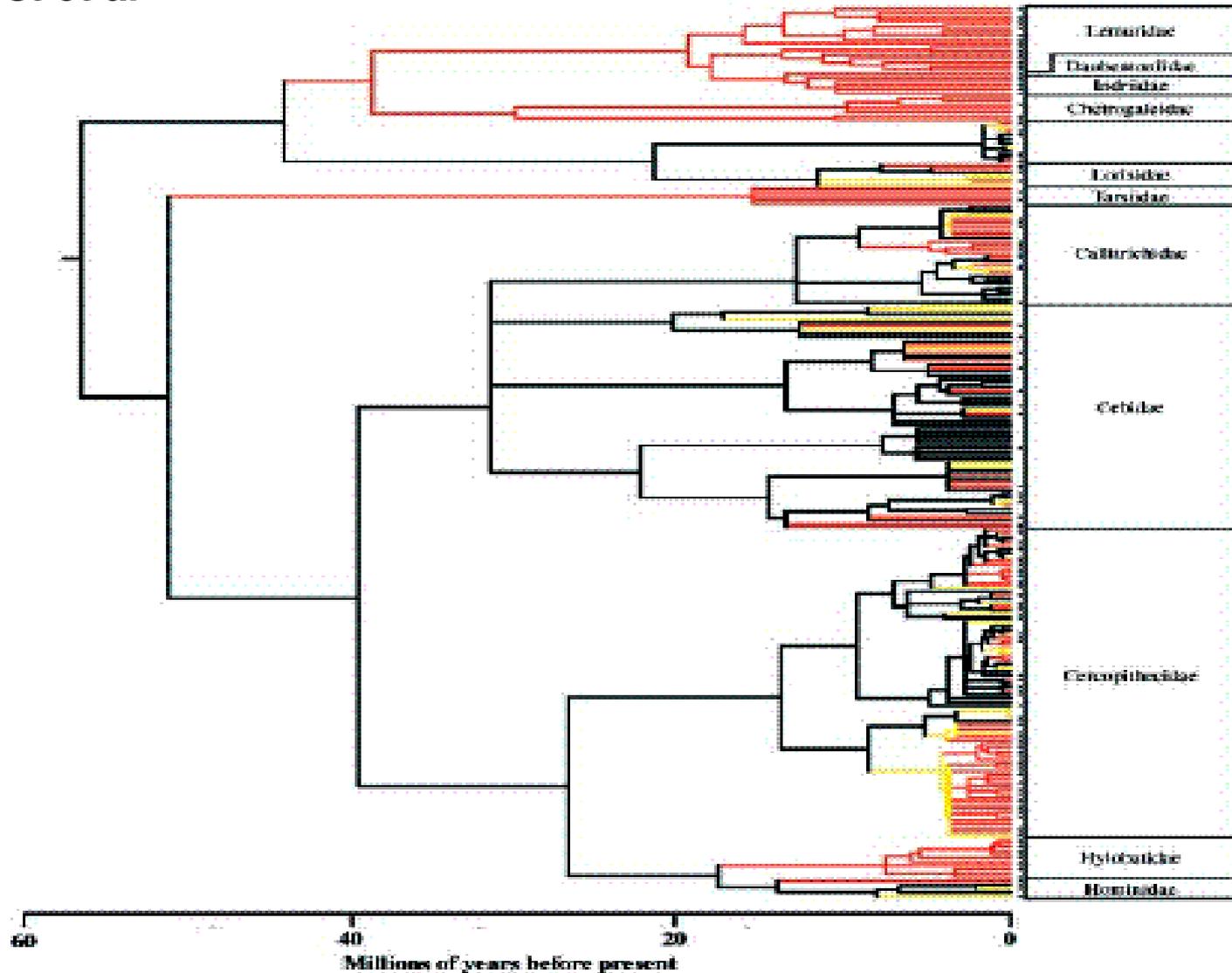
Systematic conservation planning uses PD complementarity values (gains and losses).

Planning also may use PD-endemism values

phylogenetic or PD-endemism

e.g. red branches restricted to hotspot regions

Sechrest et al



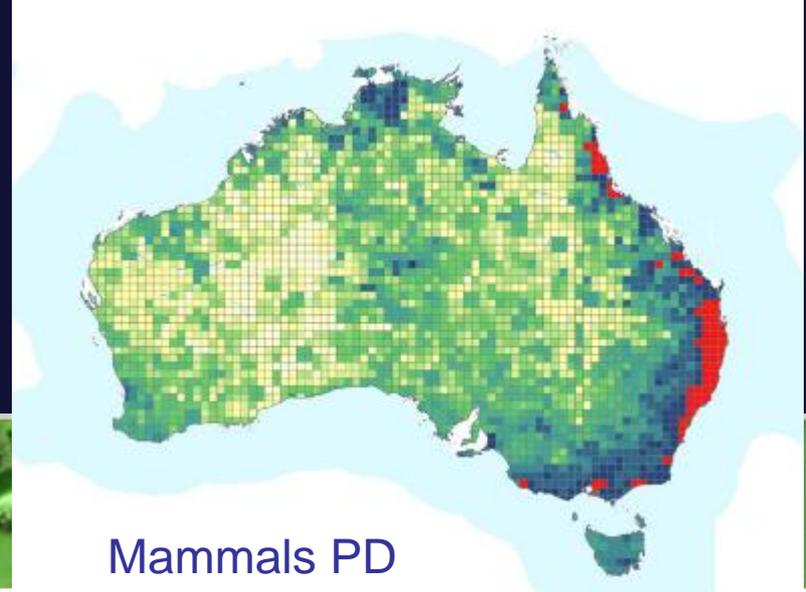
PD is a standard measure of the “evolutionary heritage” of a region or country (see for example, Mooers and Atkins 2003; Soutullo et al 2005).

However, the unique PD (endemic PD or phylogenetic endemism) of a region may have greatest heritage significance

In Australia, Faith (1994) estimated the PD restricted to Cape York for orchids, and restricted to NW Tasmania for amphipods.

Faith, D. P. 1994. Phylogenetic diversity: a general framework for the prediction of feature diversity. Pages 251-268 in: Systematics and Conservation Evaluation. Forey, P. L., Humphries, C. J. and Vane-Wright, R. I (eds). Clarendon Press, Oxford.

Phylogenetic diversity and evolutionary heritage



INTEGRATING MEASURES OF DIVERSITY AND ENDEMISM



Integrating Measures of Phylogenetic and Taxonomic Diversity and Endemism into National Conservation Assessment

Traditionally conservation planners and natural resource managers have looked to measures such as species richness, endemic species and presence and numbers of threatened species as surrogates for measuring the importance of the biodiversity in an area. Phylogenetic measures add another layer to these types of assessments and provide important information regarding the significance of the evolutionary history of an area.

The spatial database underpinning the Australian Natural Heritage Assessment Tool (ANHAT) represents an

Apply probabilistic PD – e.g. can look at the loss in expected PD if a given area is lost

Loss should be large to extent that area has long branches found in few descendants and few other areas

$\Delta PDq = \text{expected PD before loss} - \text{expected PD after loss}$

$$= \sum_i d_i \{ L_i \times (1 - q^{n_i}) \} - \sum_i d_i \{ L_i \times (1 - [q^{n_i-1}]) \}$$

where L_i = length of i th branch,

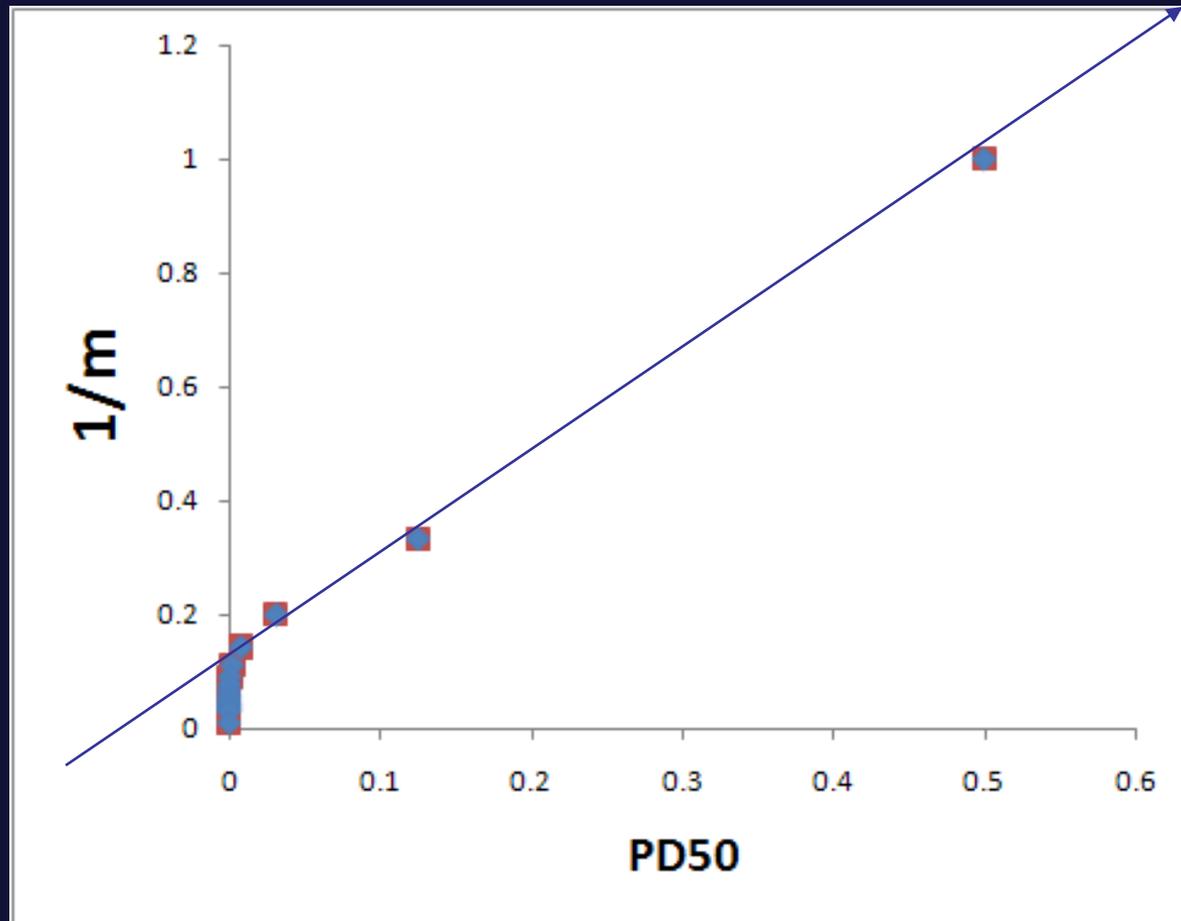
n_i = number of areas where branch descendants found,

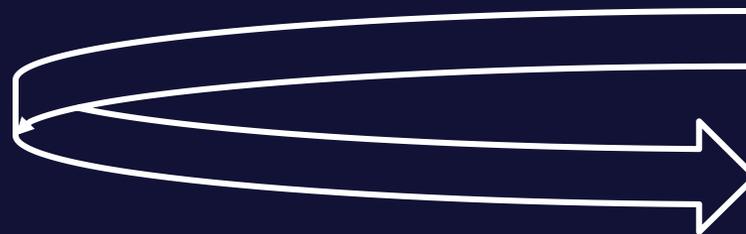
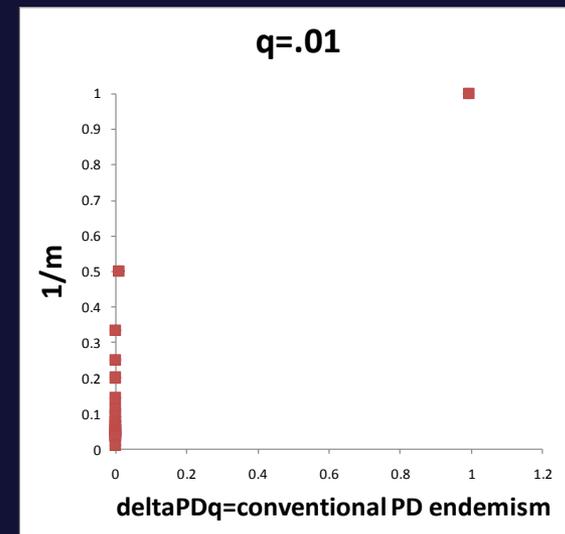
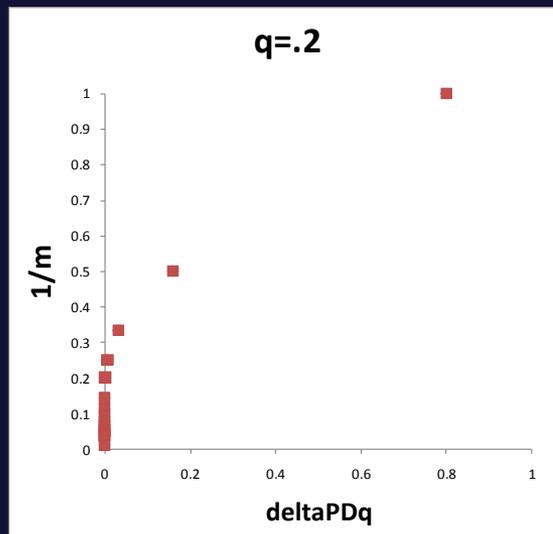
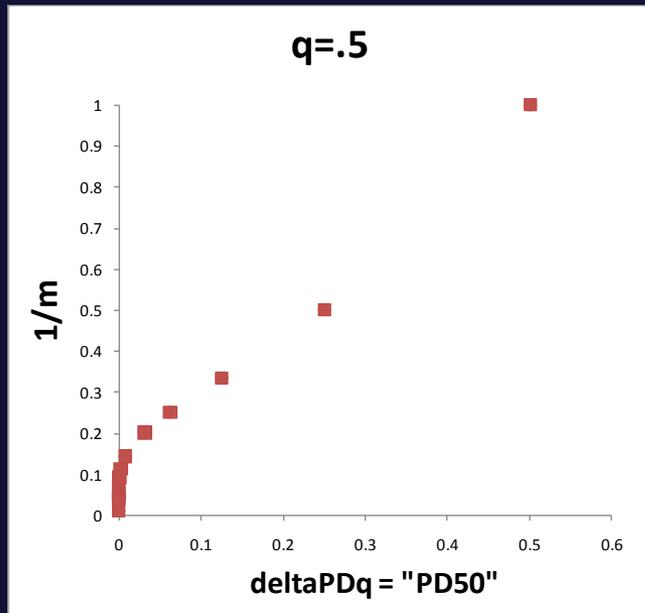
$d_i = 1$ if branch is found in the site (else 0)

and if probability of any area loss, q , = 0.5:

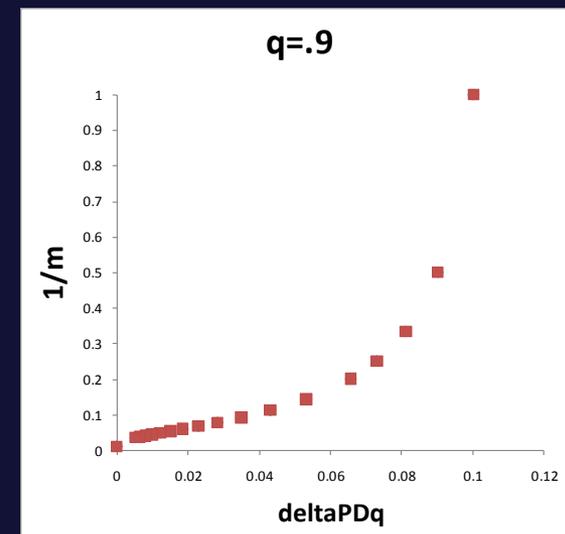
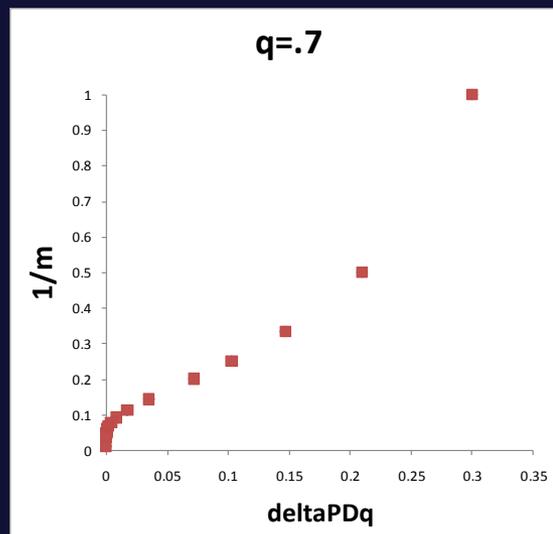
$$PD50 = \sum_i d_i (L_i \times 0.5^{n_i})$$

Compare with “PE” method where area gets score
= count of species present ,
each inverse-weighted by their total number of areas, m
The 1/m method can be over-whelmed by many
widespread species/branches in some areas.
PD50 overcomes this weakness





Greater probability,
q, of area loss



Re-examine the PD-endemism study of Faith et al 2004, for 10 taxonomic groups of beetles

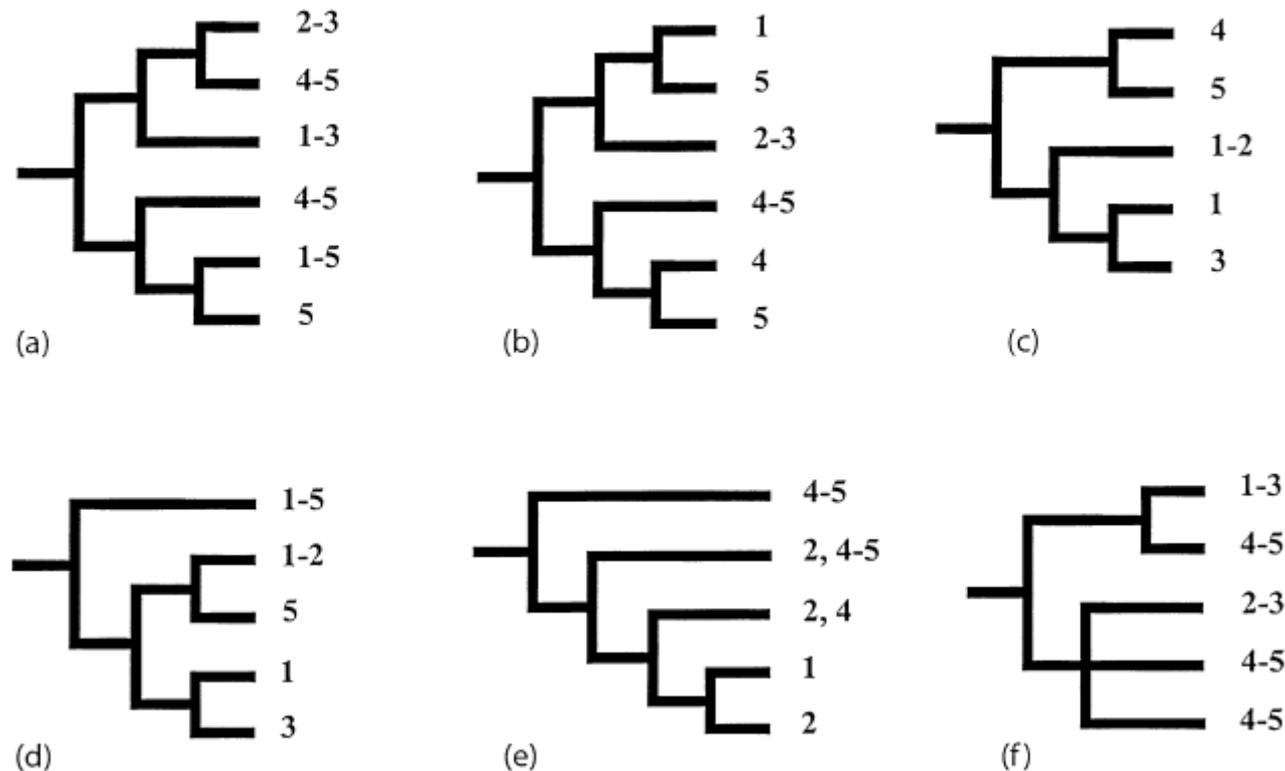


Figure 2. Estimated phylogenies for 10 taxonomic groups within Coleoptera for species in northeastern New South Wales, Australia. Branch lengths for each tree, as drawn, are arbitrary. For each phylogeny, species are labeled with numbers indicating areas of species distribution: 1, Barrington Tops massif; 2, Dorriggo Plateau; 3, eastern New England table land; 4, Mount Warning and surrounding caldera; 5, ranges around Woodenbong. The

Mount Warning, NSW (area 4) had the same PD endemism relative to area 1 (Barrington Tops). Probabilistic PD endemism (ΔPD_q) gives area 4 a higher endemism score because, in addition to having unique representation of some lineages, area 4 often shares lineages with only a small number of other areas.

Loss of area 4 means higher expected loss in PD

- it has lineages found nowhere else, and has lineages found in few other places.

Faith et al.

Phylogenetic Diversity, Complementarity, and Endemism

25

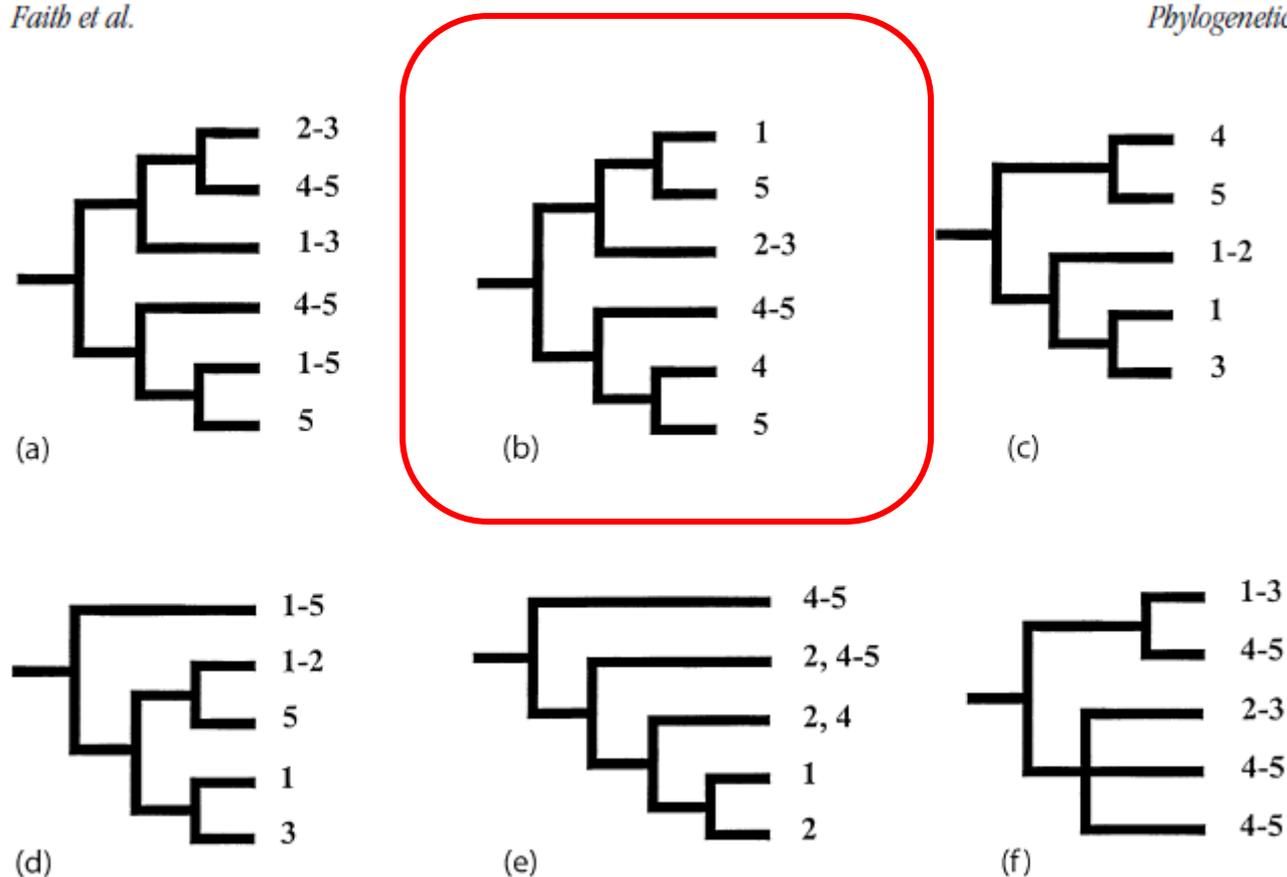


Figure 2. Estimated phylogenies for 10 taxonomic groups within Coleoptera for species in northeastern New South Wales, Australia. Branch lengths for each tree, as drawn, are arbitrary. For each phylogeny, species are labeled with numbers indicating areas of species distribution: 1, Barrington Tops massif; 2, Dorrigo Plateau; 3, eastern New England table land, Ebor Plateau; 4, Mount Warning and surrounding caldera; 5, ranges around Woodenbong. The

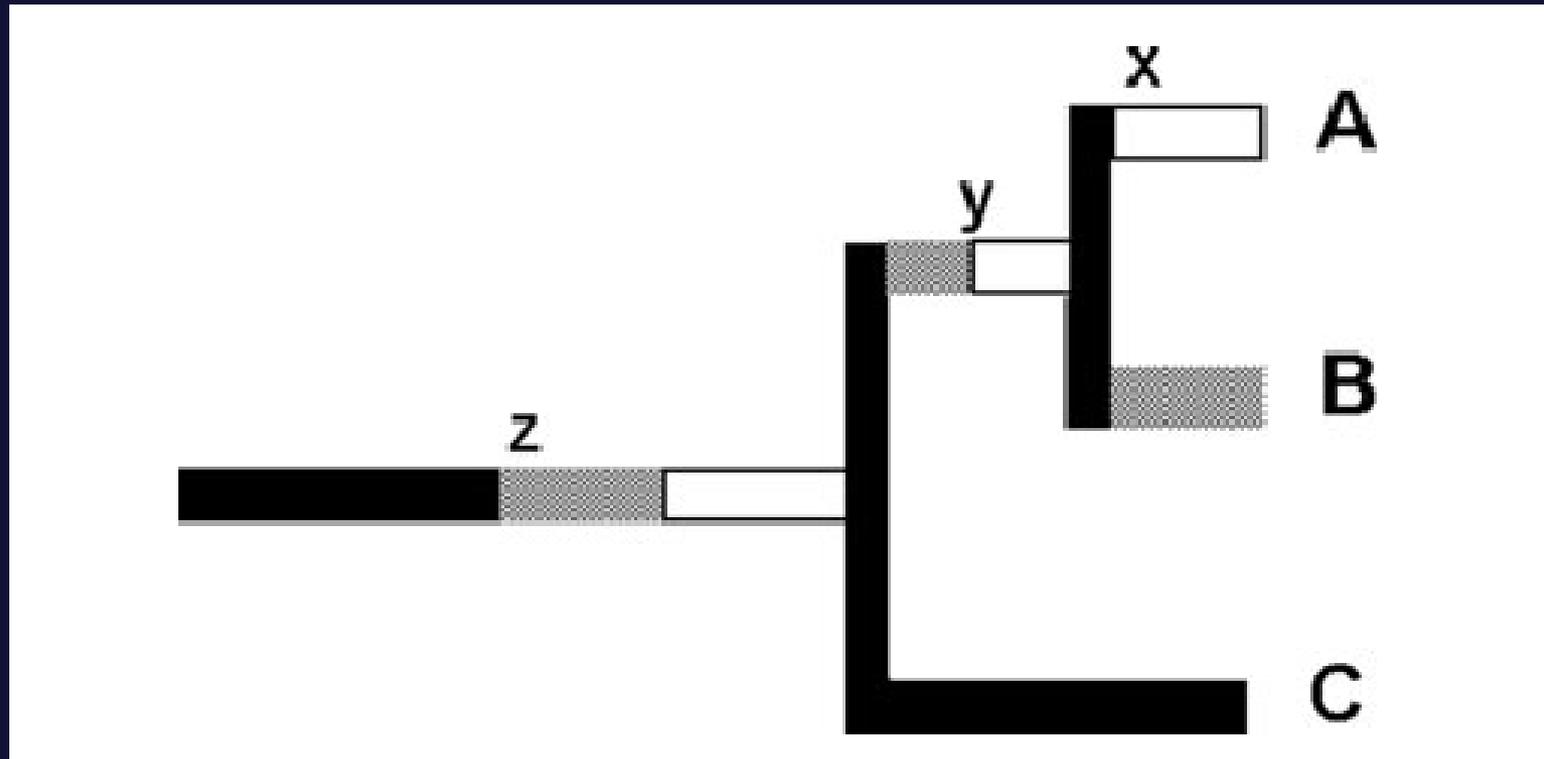
Probabilistic PD calculations therefore provide a range of indices for prioritising species and areas, and providing indices for grid cells for mapping

Contrast with a family of alternative phylogenetic calculations that do not work very well....

“evolutionary distinctiveness”
measures

A weakness of EDGE calculations (and W-type indices for phylogenetic endemism measures)

species are given scores reflecting
shared credit for deeper branches



A weakness is that scores for individual species do not sensibly
combine with probabilities of extinction,
because the degree of phylogenetic overlap is not taken into account

BED is based on the ED method underlying EDGE approaches. It divides up credit for representing branches among the species on the phylogenetic tree. Areas then receive scores based on the sum of species scores. BED extends this idea by partitioning the credit also among (for example) the grid cells used by each species. The key problem is that when species are combined to provide areas scores, BED does not take phylogenetic complementarity among species into account. Here BED would not detect cell 2 as the better choice.

a b c d x

A X X X X

B X X X X

C X X X X

D X X X X

L

L

L

L

L

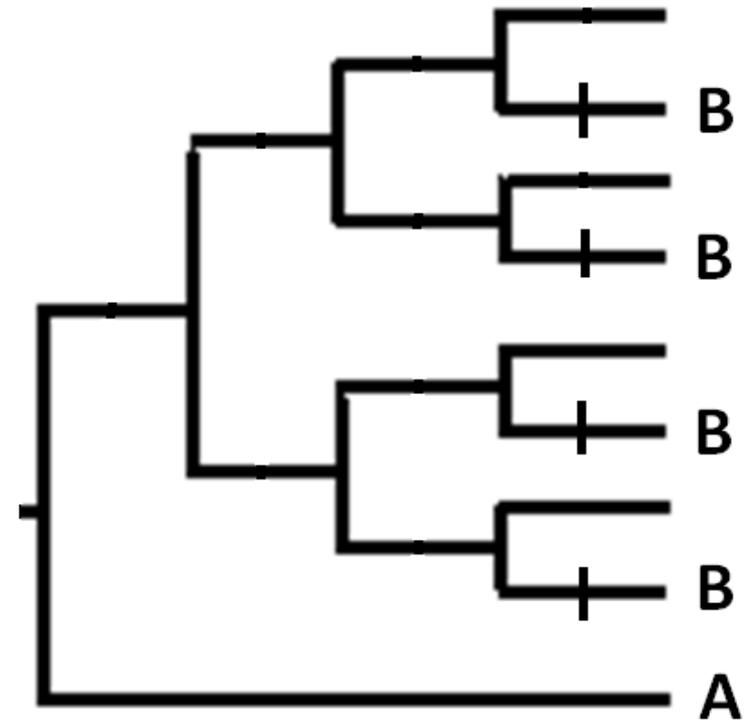
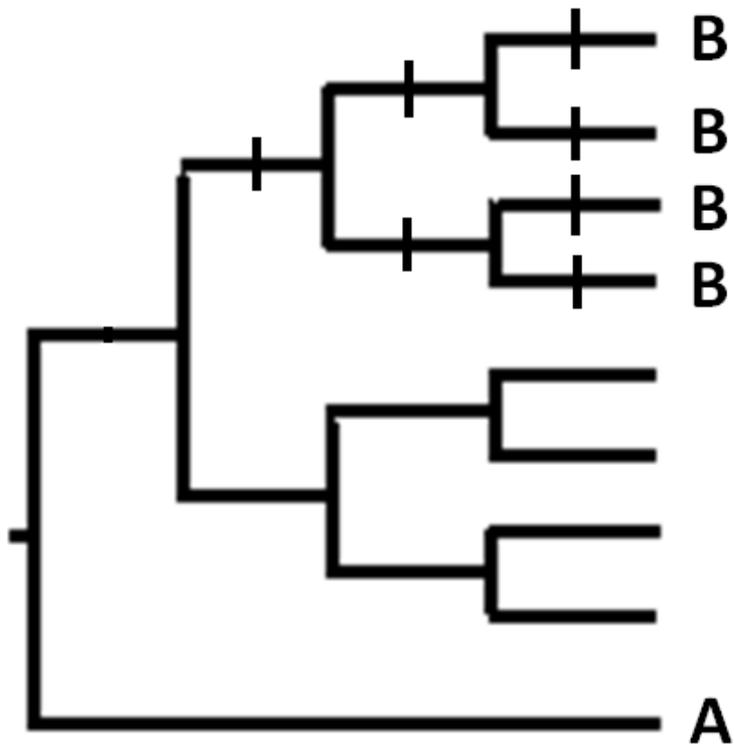
cell (1)

cell (2)

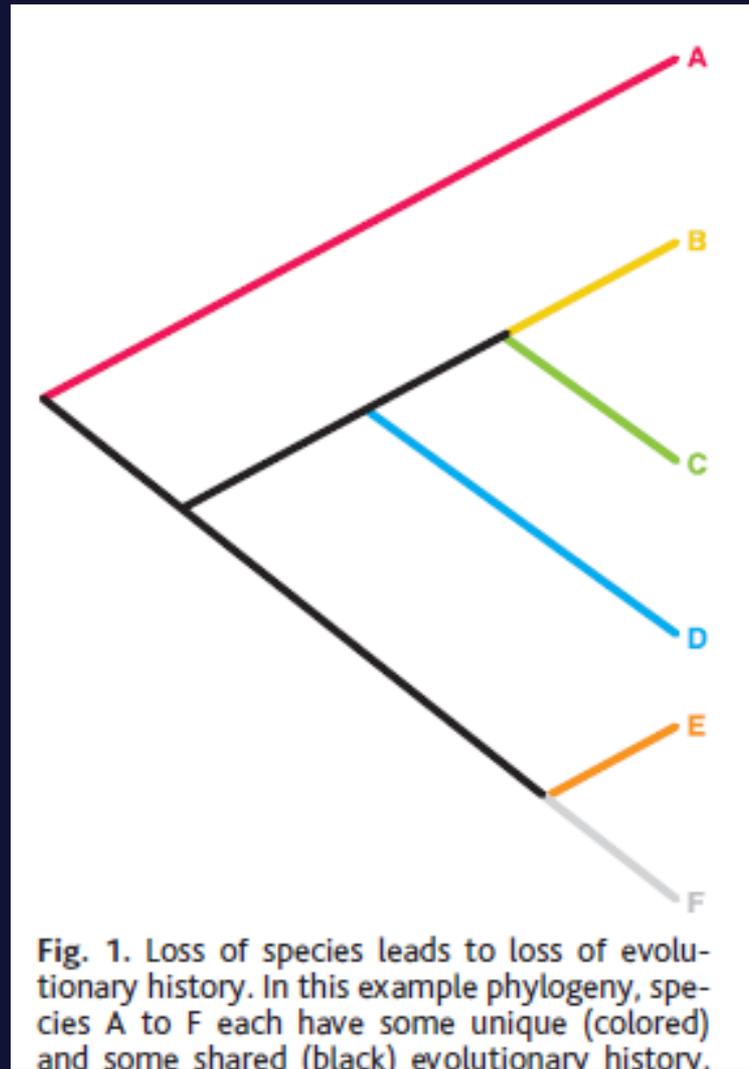
Cell (1) uniquely has species a,b,c,d, which are on small branches of length m , and are at the end of a long branch of length L . species x uniquely occurs in some other site/cell.

Cell (2) uniquely has species A,B,C,D, which are on small branches of length m , and are at end of different long branches of length L . For each member species, 4 other sister species marked X all uniquely occur in some other site/cell.

A weakness is that scores for individual species do not sensibly combine to give scores for areas because the degree of phylogenetic overlap is not taken into account



The key concept of shared evolutionary history also leads to a simple definition of PD-dissimilarities among samples or areas (“phylogenetic beta diversity” sensu Lozupone and co-workers)

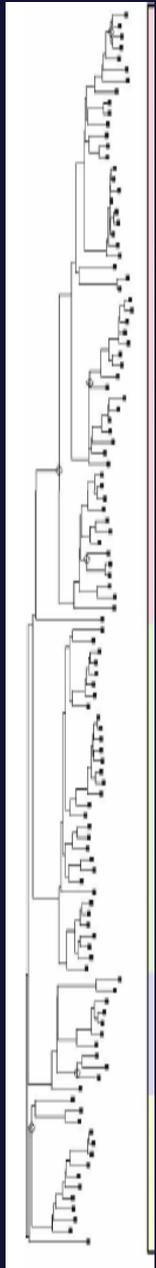


Mace et al (2003); Science 300, 1707

Microbial ecology

two sample sites j and k

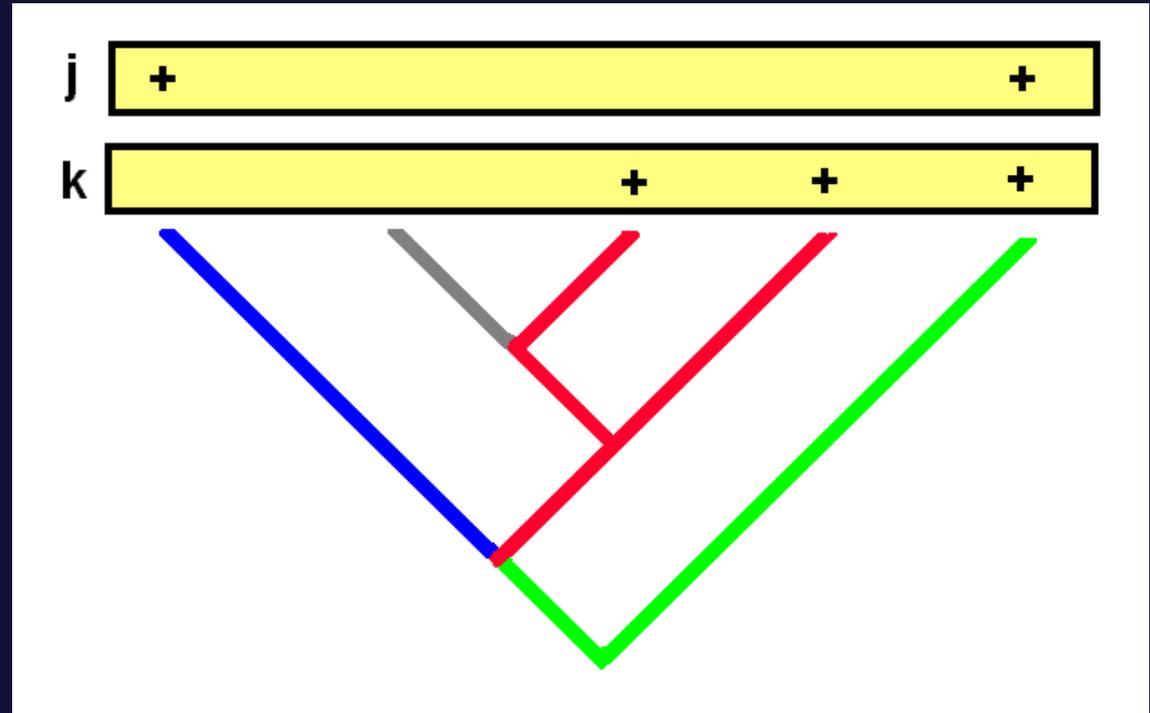
j and k dissimilar if lots of red and blue



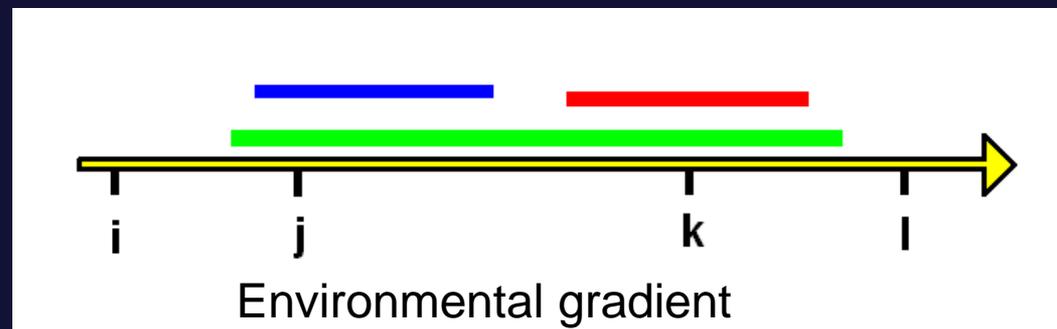
Big trees,

Few taxa,

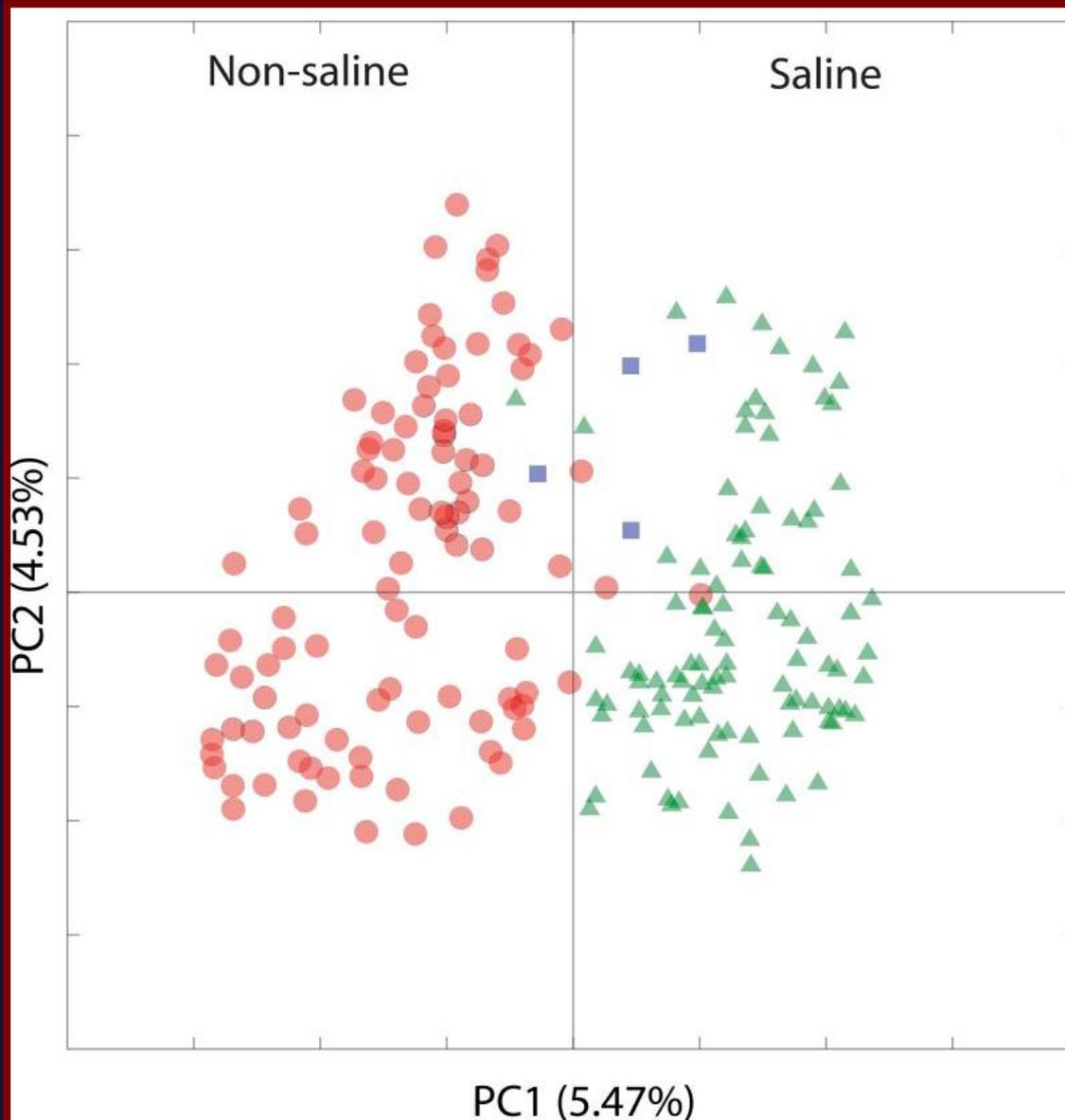
Many samples



PD-dissimilarities reflect distances along gradients

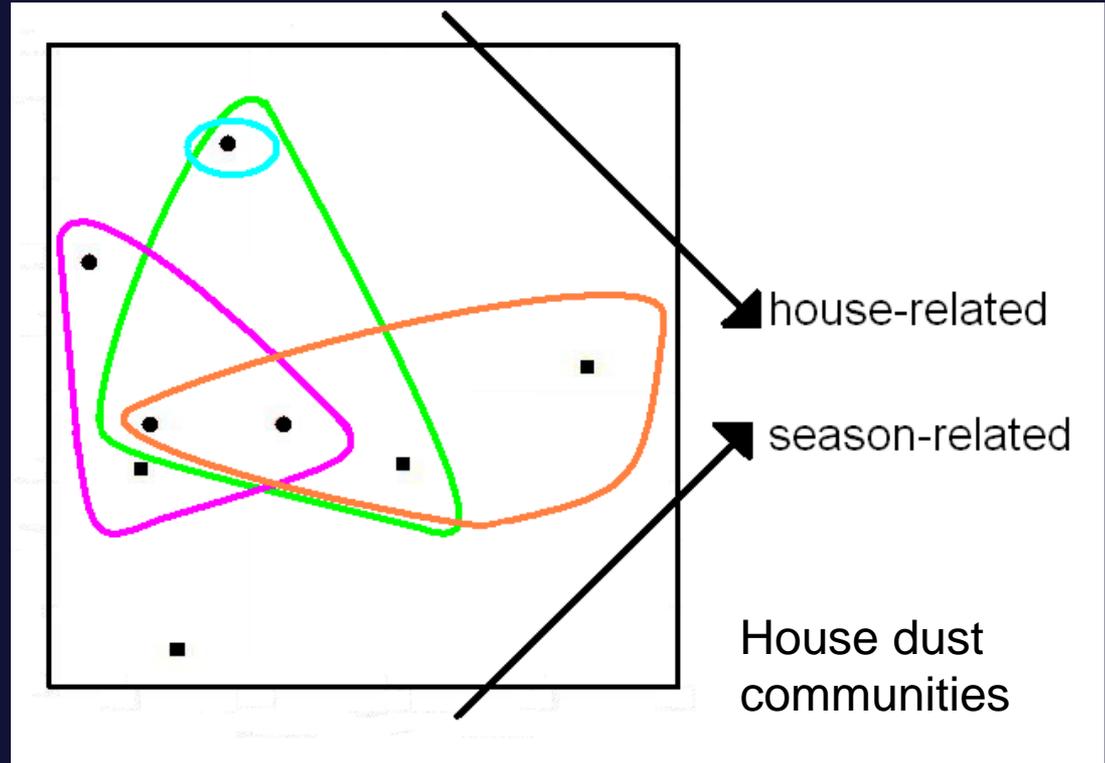
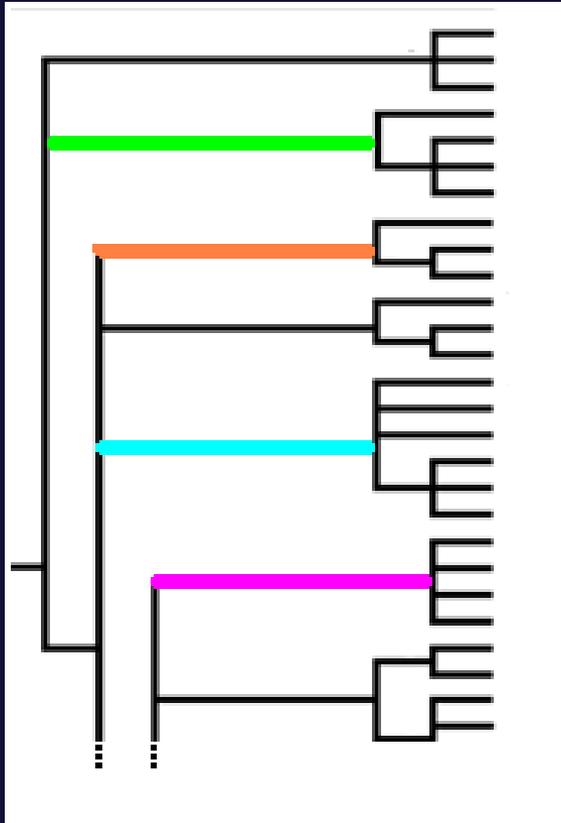


Lozupone and Knight 's “phylogenetic beta diversity” for global bacteria samples



- Use phylogenetic dissimilarity (“UniFrac”) among samples
- Discover that the major environmental determinant of microbial community composition is salinity

Phylogeny helps find important gradients, because even deeper branches have unimodal response to gradients



Faith, D. P., C. A. Lozupone, D. Nipperess, R. Knight
A general model linking evolutionary features and environmental
gradients supports broad applications of microbial ecology's phylogenetic
beta diversity framework. *International Journal of Molecular Science*

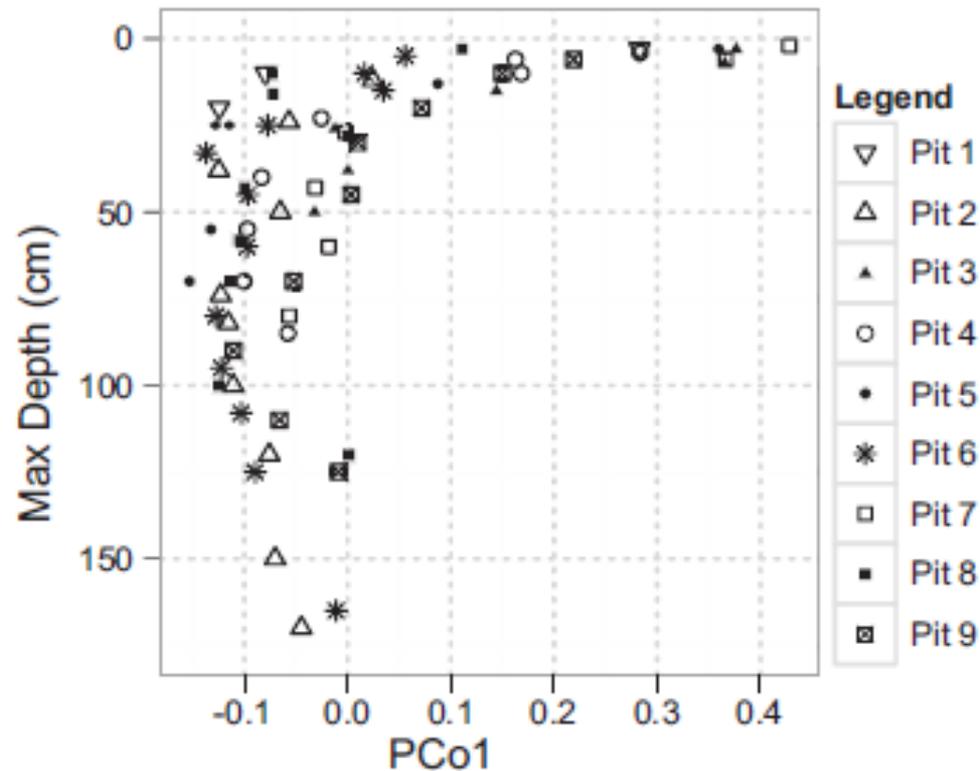


Fig. 3. Change in microbial community composition with depth (beta diversity) as measured by the score on the first principle coordinate axis of a principle coordinates analysis of the weighted UniFrac distance matrix.

Common weakness of ordination analyses of PD-dissimilarities - use PCA and other methods that are not compatible with “unimodal” response of features/lineages to environmental gradients

Phylogeny and functional trait diversity

Traits-based measures typically focus on a nominated set of “important” traits.

e.g. for the FD method, Petchey & Gaston (2002) argue “the measured traits should be those for which evidence exists of their functional importance”

Safi et al used Gower distance and the unweighted pair group method with arithmetic averages to produce the distance matrix and the functional dendrogram – then applied PD calculations

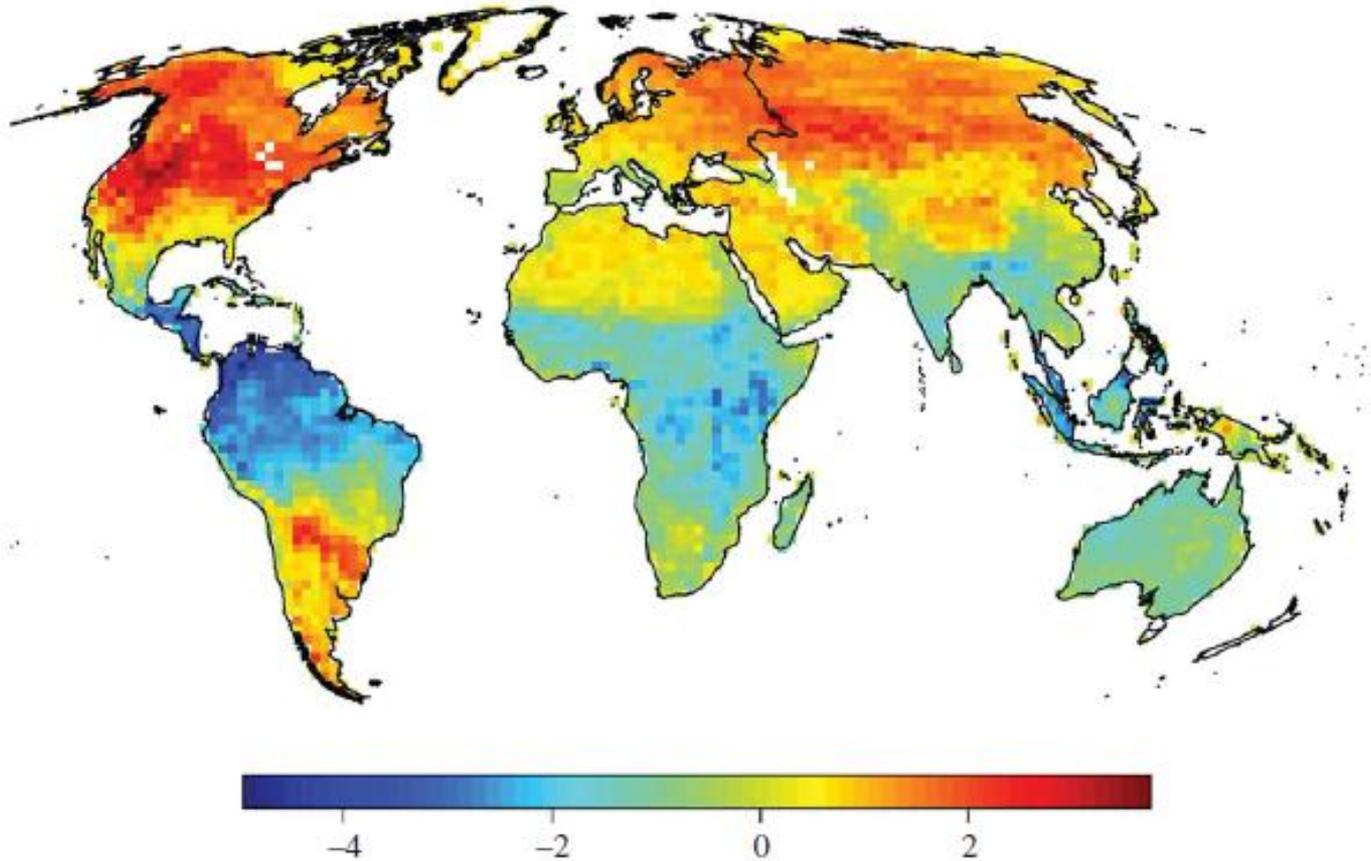
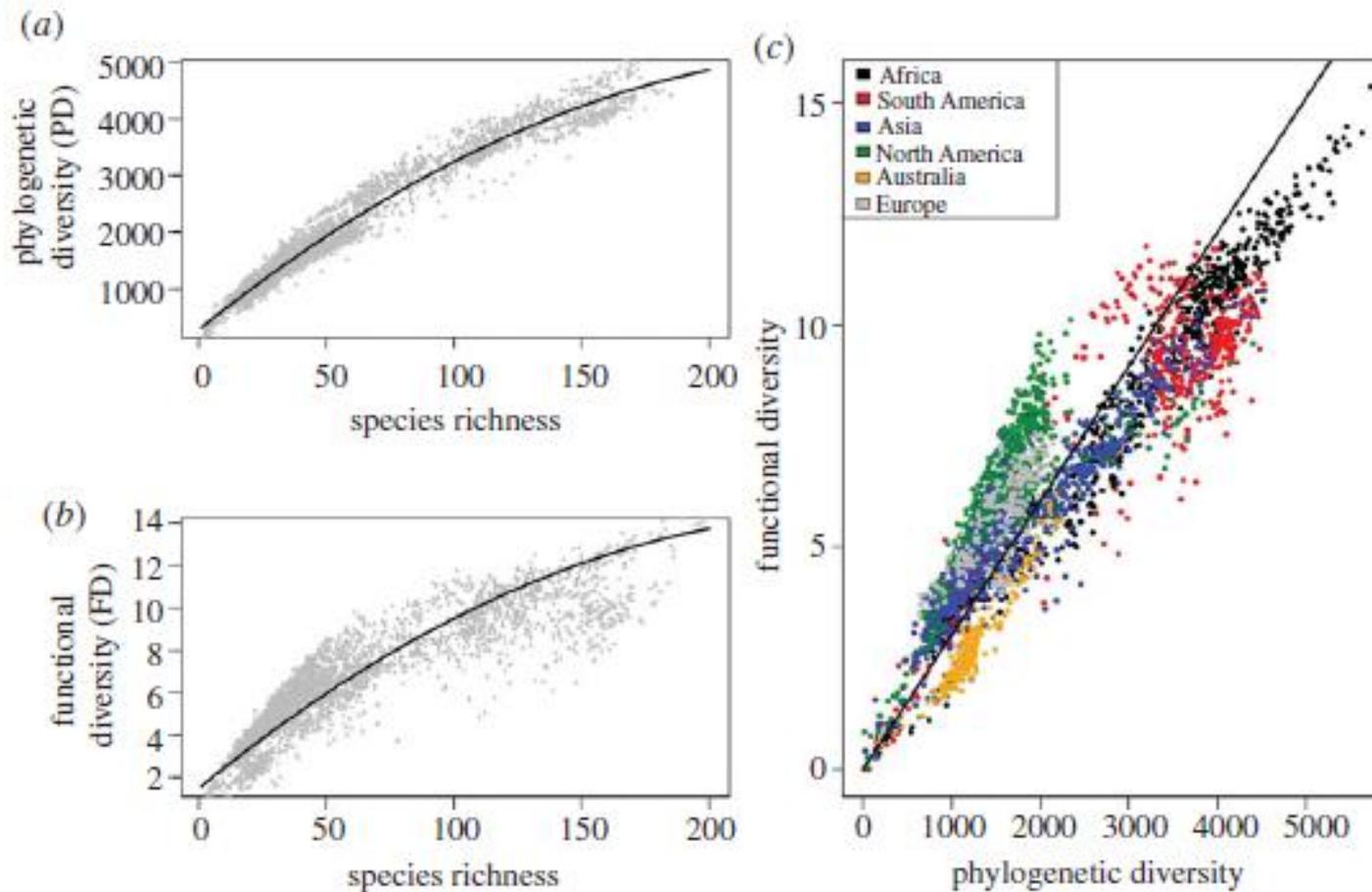


Figure 3. Deviance (residuals) of observed FD compared to the global linear relationship between PD and FD ($FD = 0.003 \cdot PD$; see figure 2c). Blue areas depict areas of lower than expected FD (according to the PD present in the area) whereas red areas are areas with more FD than expected from this linear relationship.

Safi et al (2011) *Phil. Trans. R. Soc. B* 366, 2536-2544.

Understanding global patterns of mammalian functional and phylogenetic diversity



Safi *et al.* (2011) *Phil. Trans. R. Soc. B* 366, 2536-2544.

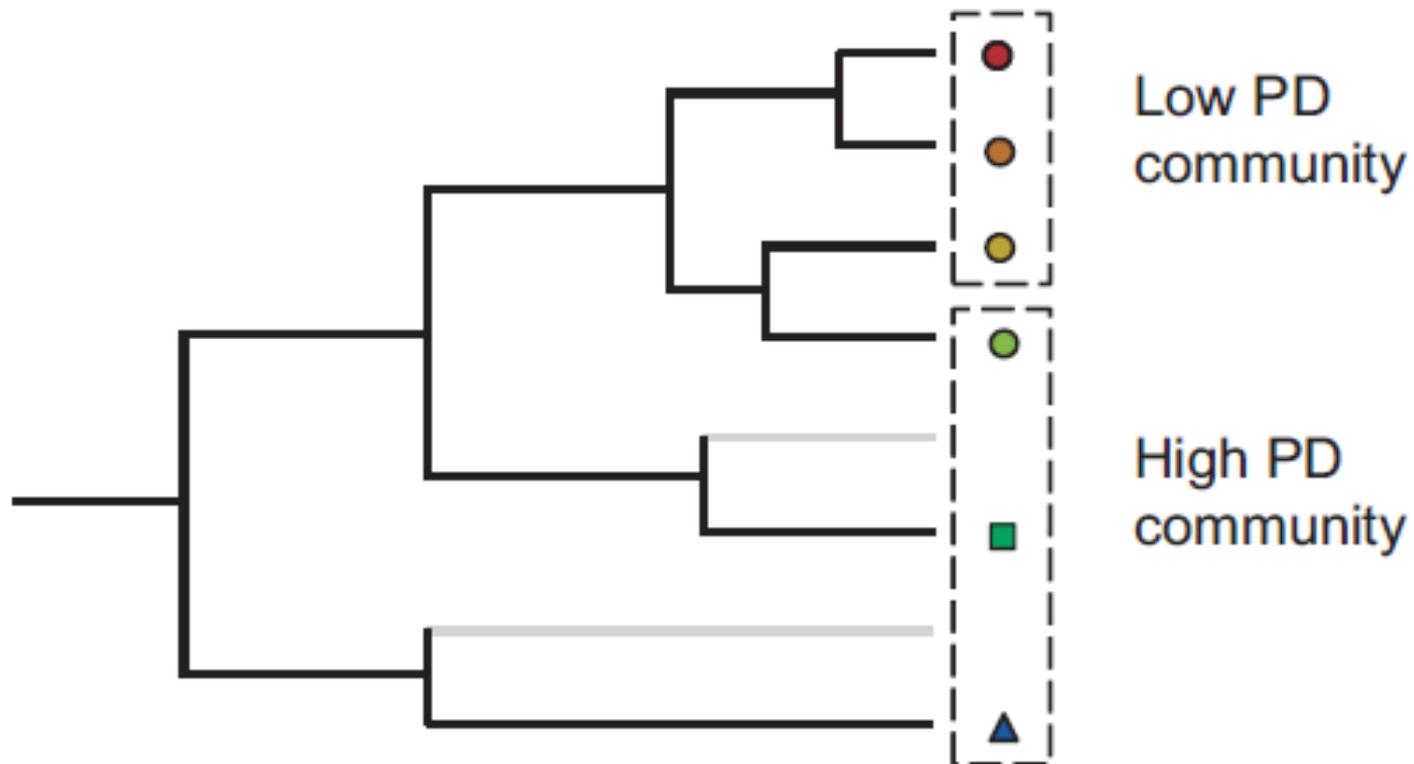
Understanding global patterns of mammalian functional and phylogenetic diversity

Need measures of general functional trait diversity.....

Some evidence for PD's proxy-value for functional diversity

“Many traits show a phylogenetic signal, suggesting that PD can estimate the functional trait space of a community, and thus ecosystem functioning”

(a) Phylogenetic signal exists for functional traits

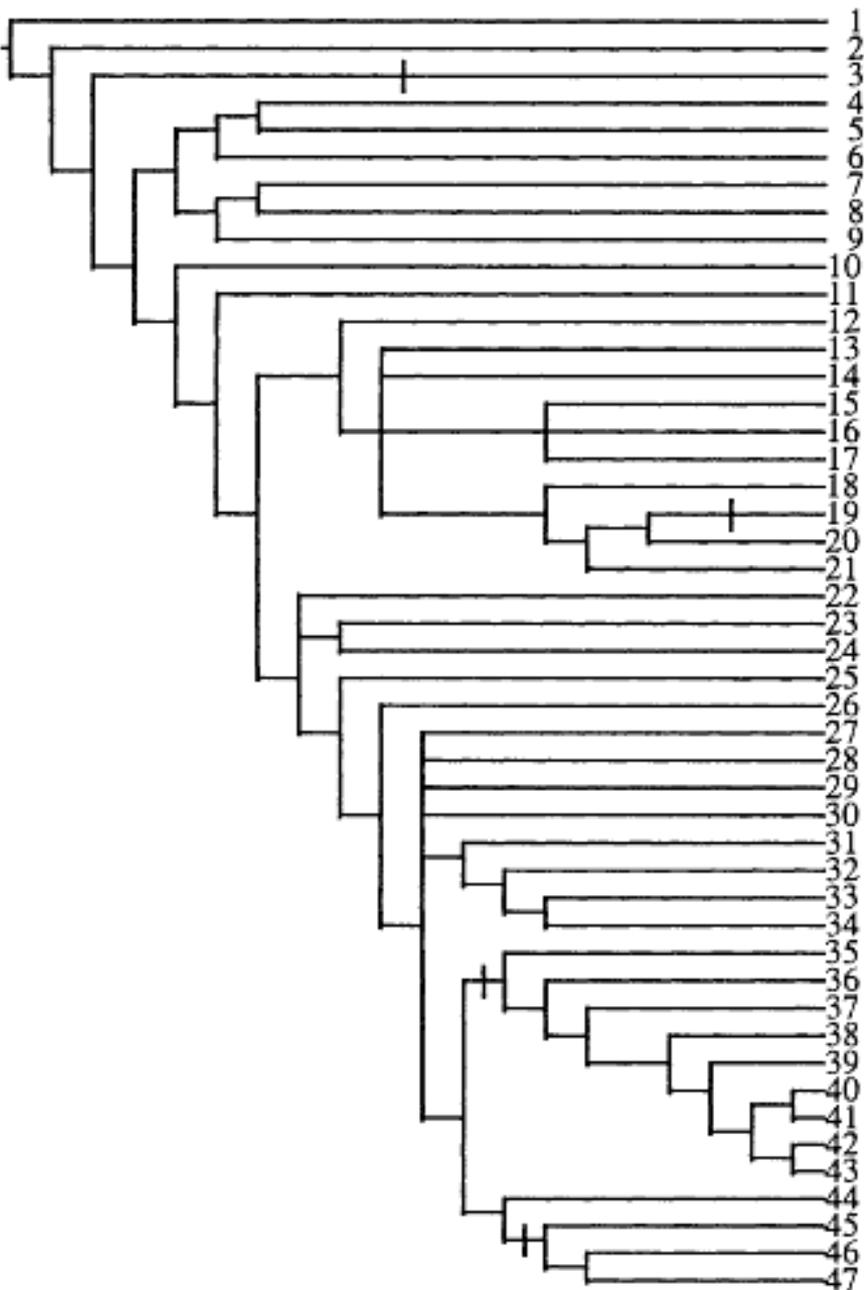


Some evidence for PD's proxy-value for functional diversity but limitations recognised. Weiher et al. (2011) observe that high PD may not correspond to high trait diversity because of convergent evolution.

In fact, long ago there was a suggested alternative to PD to address functional trait diversity.

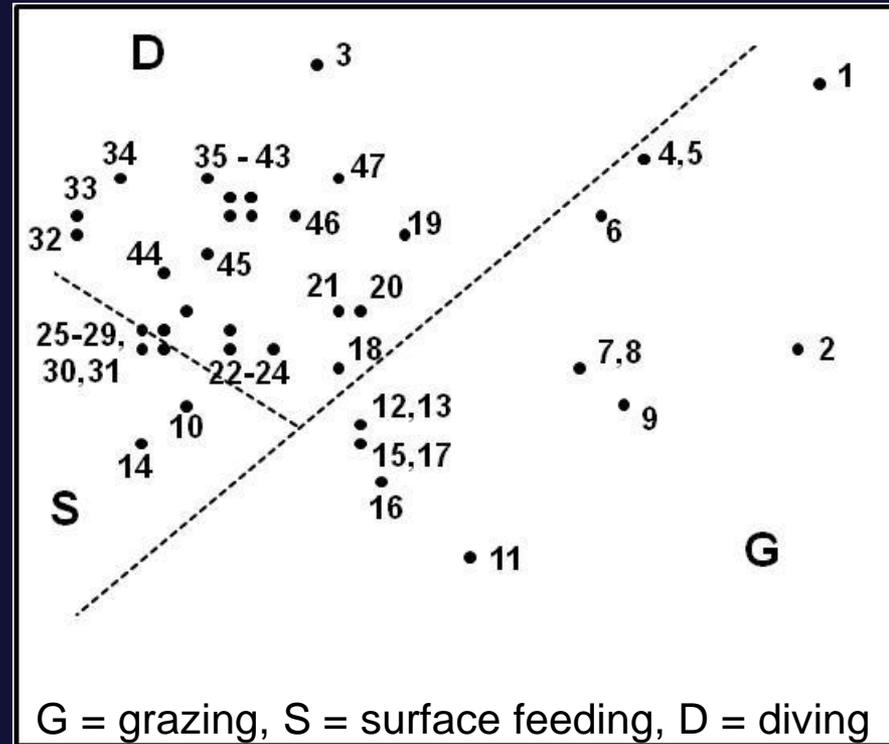
This measure considers convergent evolution.

Faith 1996 *Conservation Biology*

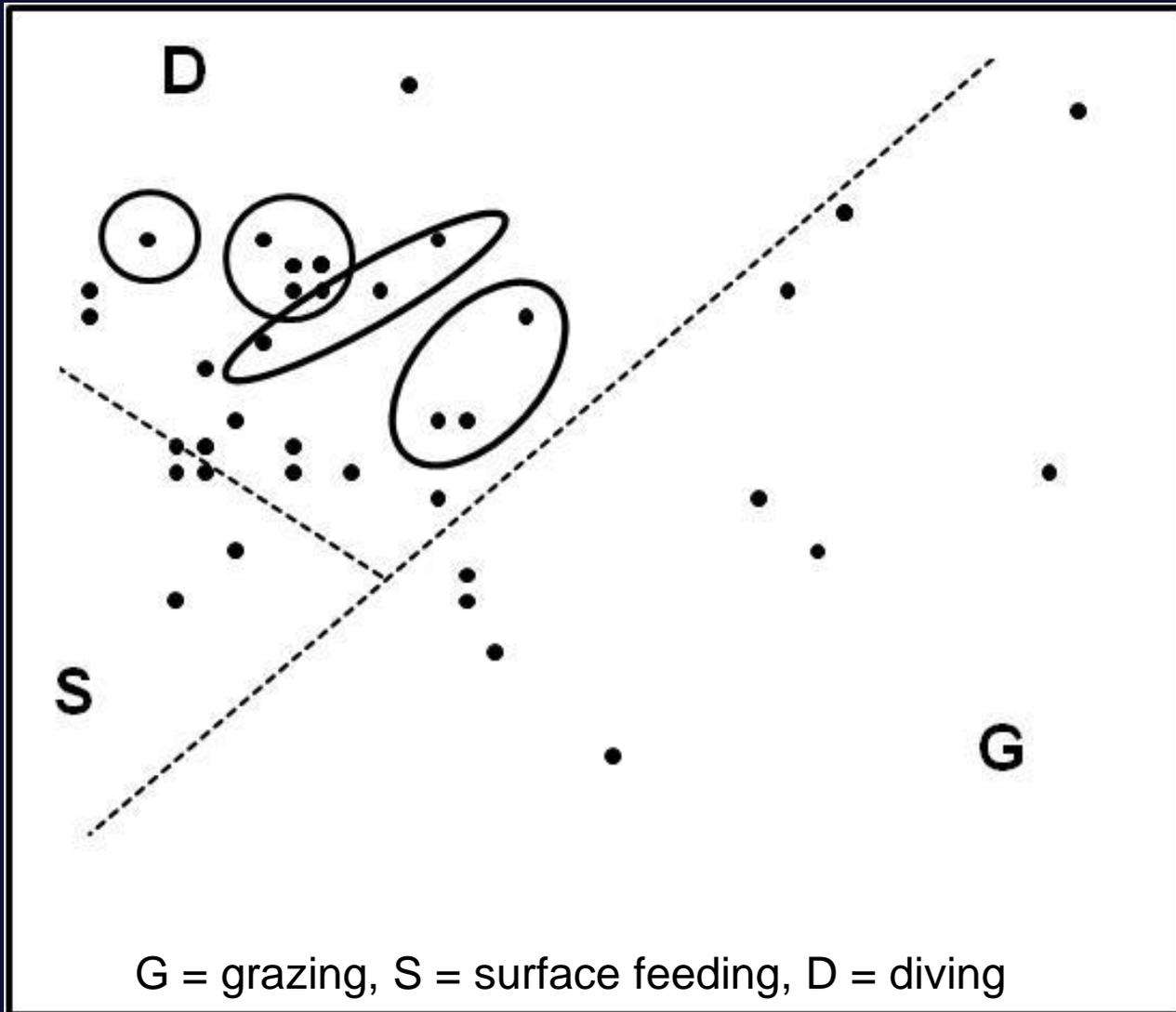


For phylogenetic tree for Anseriformes, analyse data matrix made up only of the convergently derived characters

Get a pattern where taxa close together may not be closely related phylogenetically



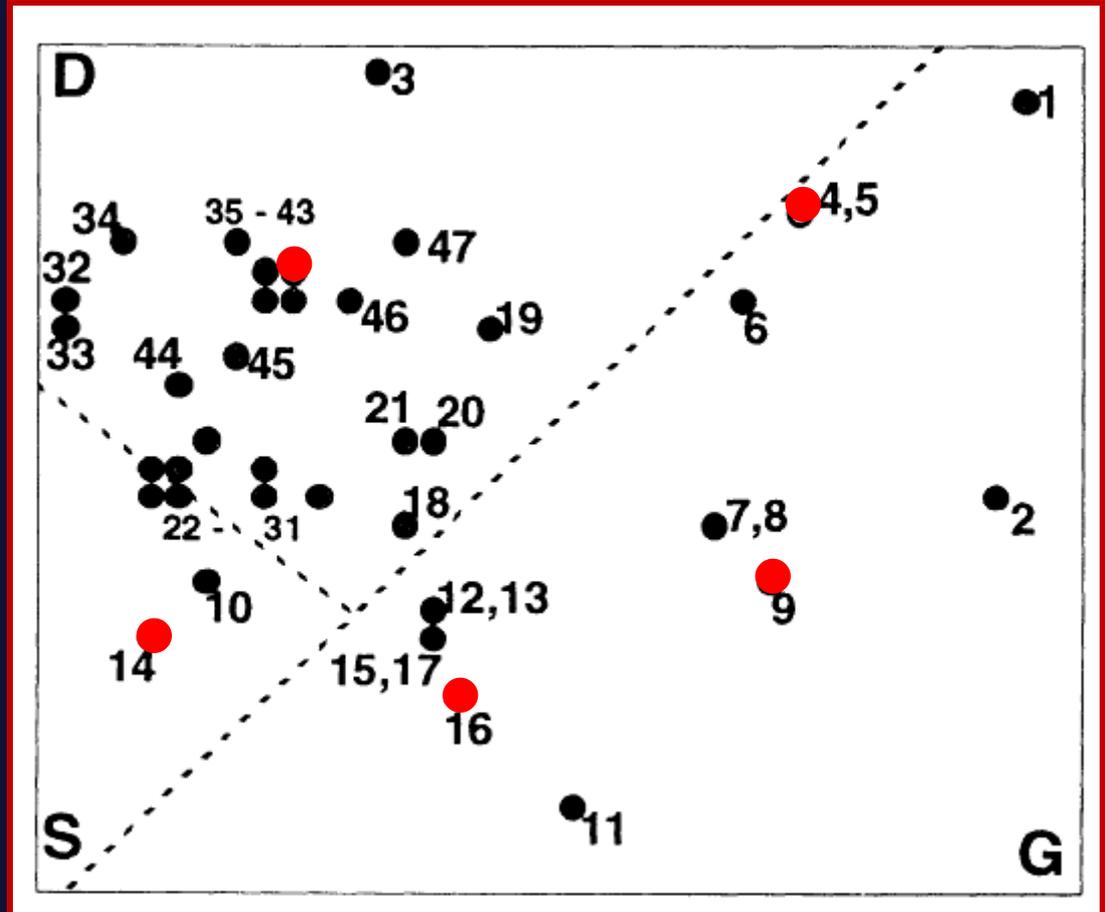
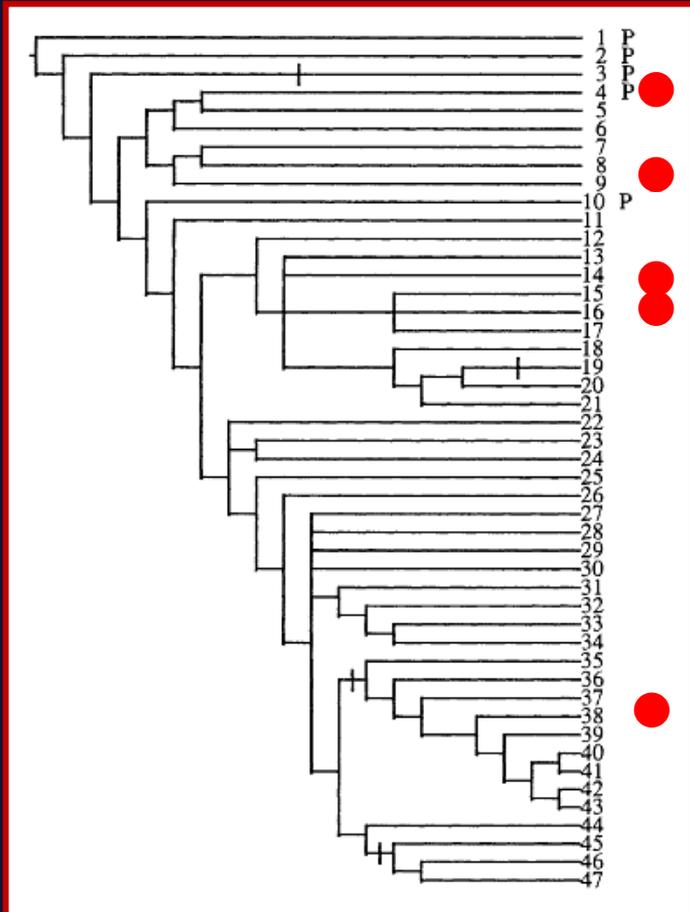
Example of one convergently derived feature.
For details see Faith 1989;1996



phylogeny and functional diversity

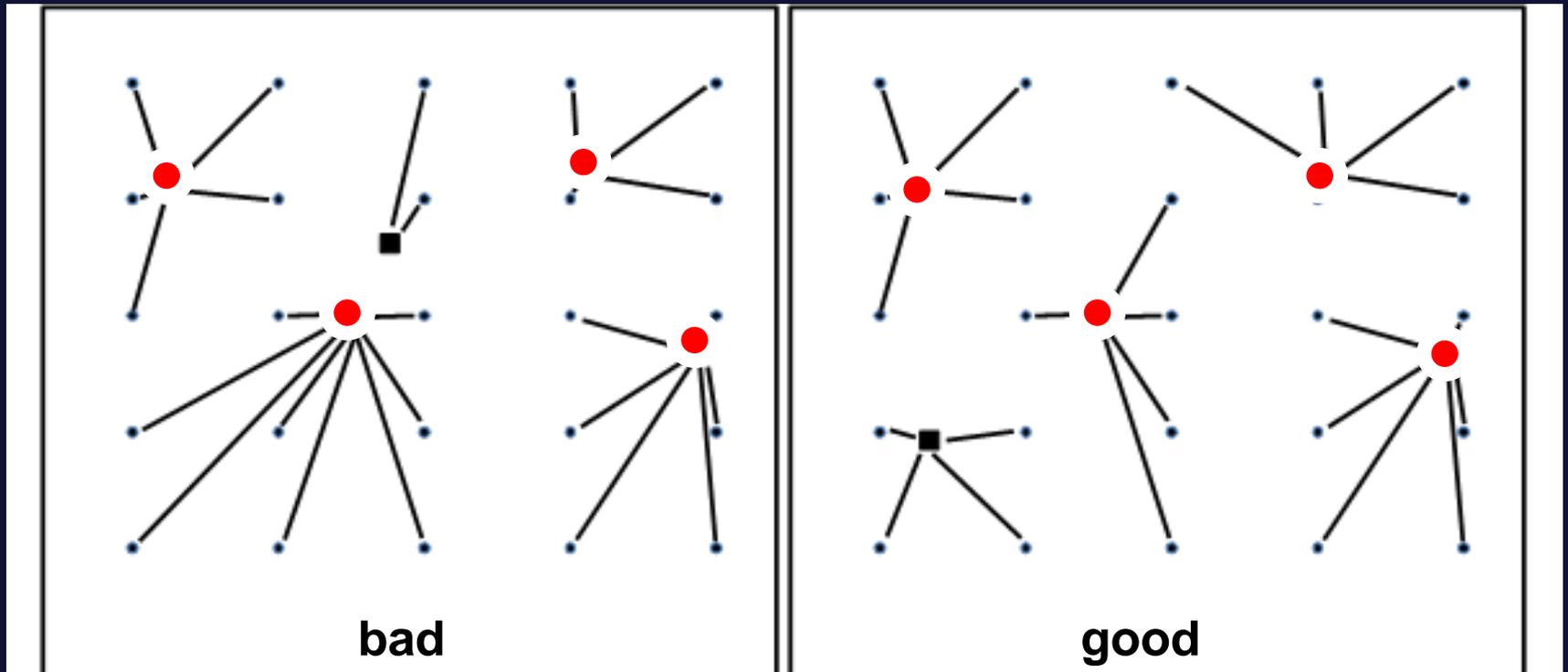
In tree on left, "P"s mark best 5 species for PD.

In functional space at right, red dots mark best 5 species.



G = grazing, S = surface feeding, D = diving

A method called ED ('environmental diversity') provides a measure of the functional trait diversity of any subset of taxa



questions