

A range of phylogenetic tools and methods
for biodiversity conservation
based on the PD
(phylogenetic diversity) measure

Daniel P Faith

The Australian Museum

Sydney, Australia

Hennig XXX - 2011 Meeting of the Willi Hennig Society
Hotel Nacional Plaza Inn, São José do Rio Preto,
July 30 - August 2

bioGENESIS

Providing an
evolutionary framework
for biodiversity science



BioGENESIS

Loss of evolutionary history is a major concern within the broader biodiversity crisis.

Recently highlighted as the loss of “**eVosystem services**”

Faith, D.P., S. Magallón, A.P. Hendry, E. Conti, T. Yahara, and M.J. Donoghue. (2010). Evosystem Services: an evolutionary perspective on the links between biodiversity and human-well-being. *Current Opinion in Environmental Sustainability*.

GEO BON: a global Biodiversity Observation Network



Key challenges for GEO BON include finding effective ways to better observe and monitor at the level of genetic and phylogenetic variation.

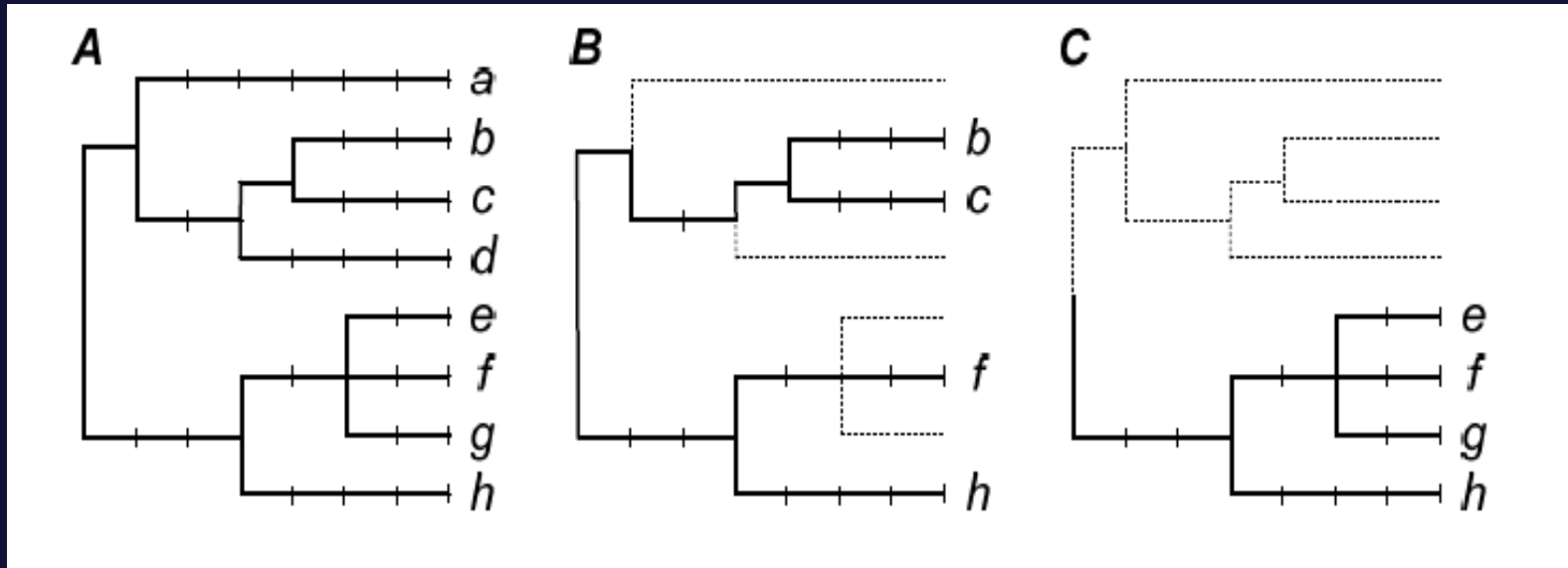
Two strategies for monitoring phylogenetic diversity

Repeated observations, over time, of:

- 1) Specific phylogenetic components of interest, in selected target groups.
- 2) changes in land/water condition (e.g. using remote sensing), integrated with spatial phylogenetic variation models as the “lens” to infer corresponding changes at the phylogenetic levels.

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”

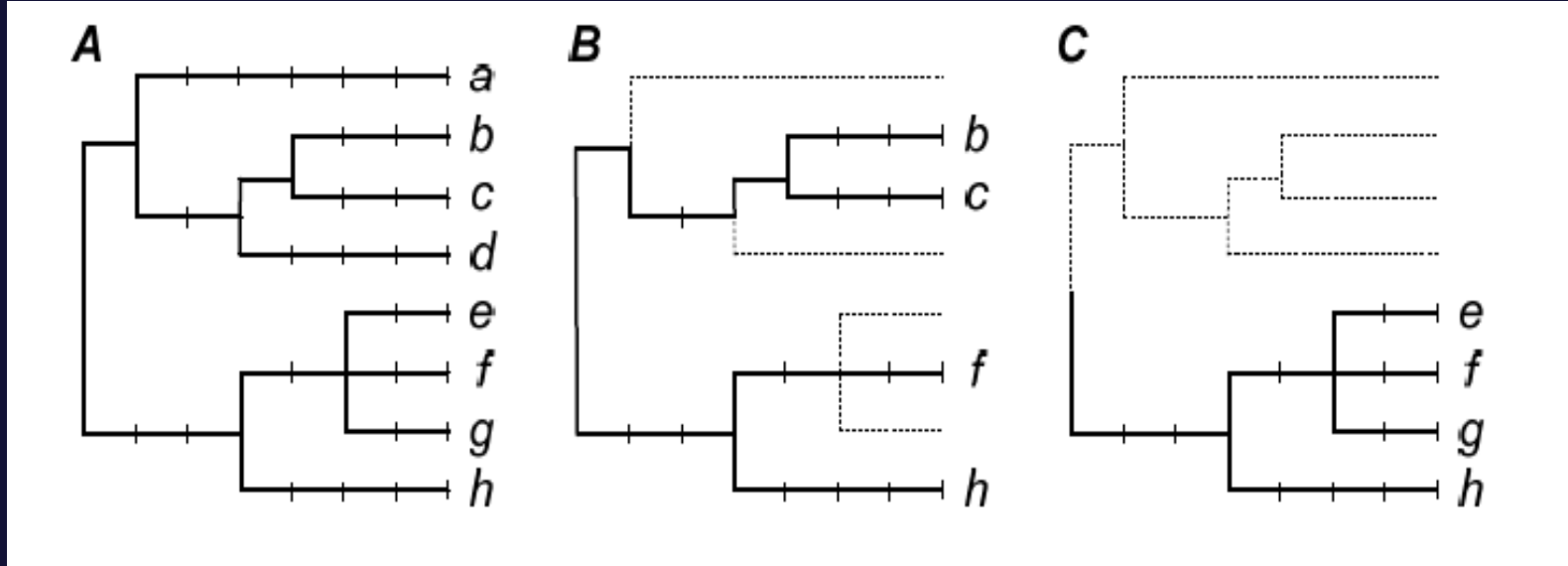
scenario *B* represents more feature diversity

Faith DP. *Biological Conservation* (1992).

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

PD – phylogenetic diversity

Faith 1992



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

Richness = total PD

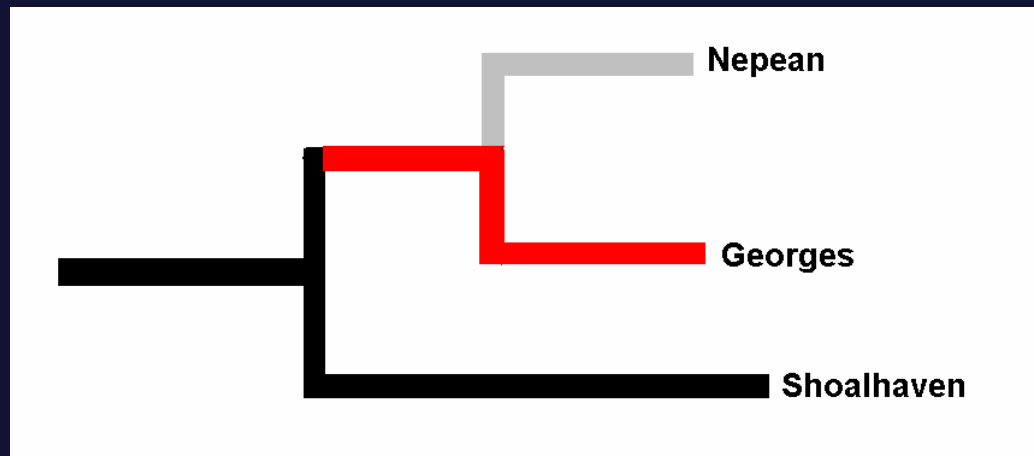
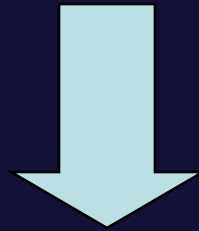
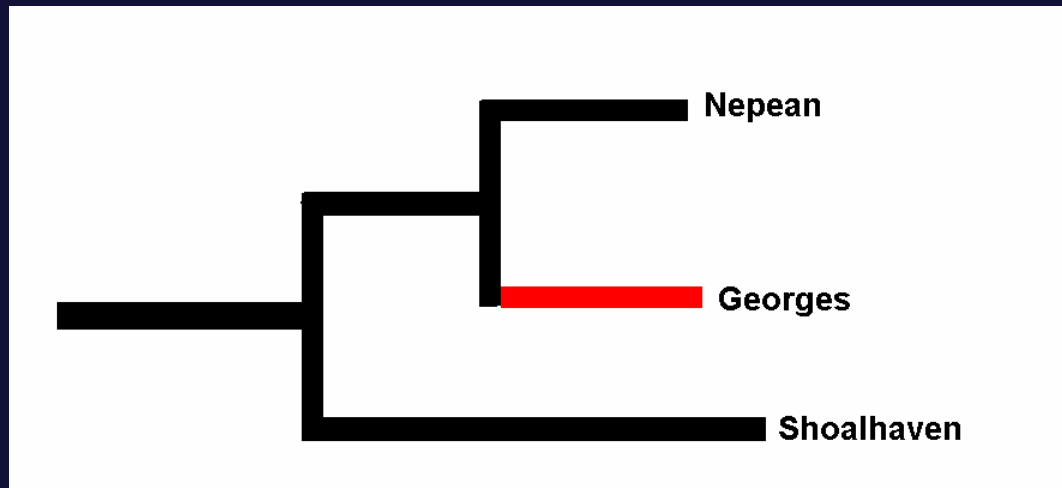
Expected diversity = expected PD

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



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.

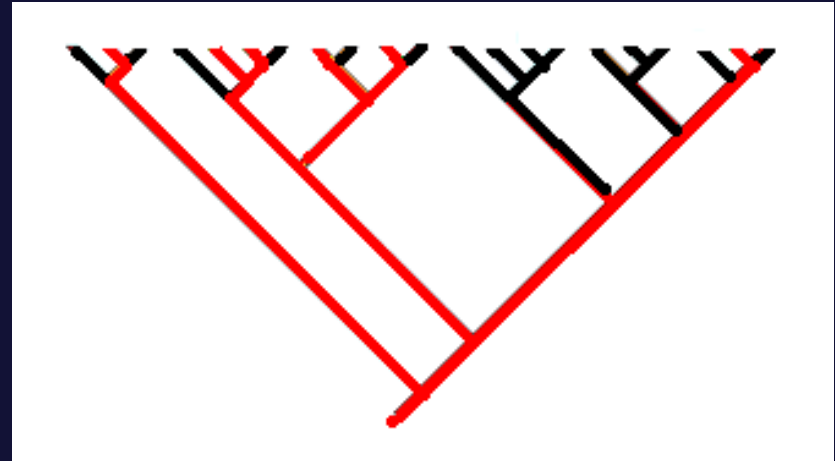
Will the impacts of climate change on PD be large or small?

Yesson, C. and A. Culham. 2006.

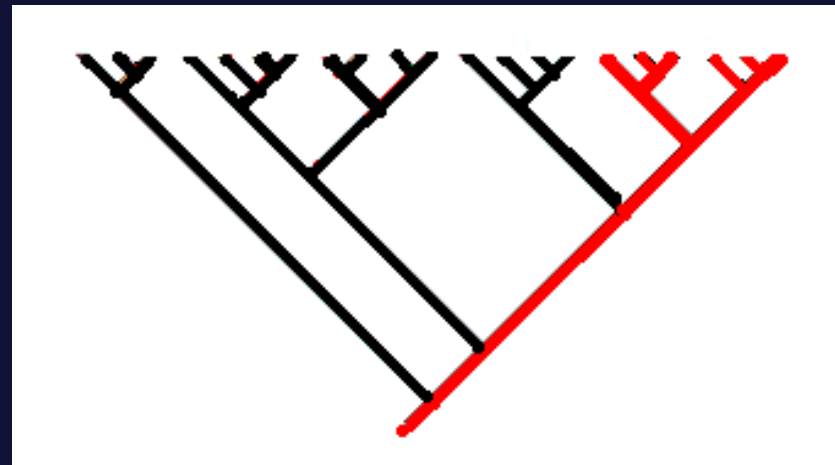
- *small* loss of PD or evolutionary potential for given species loss



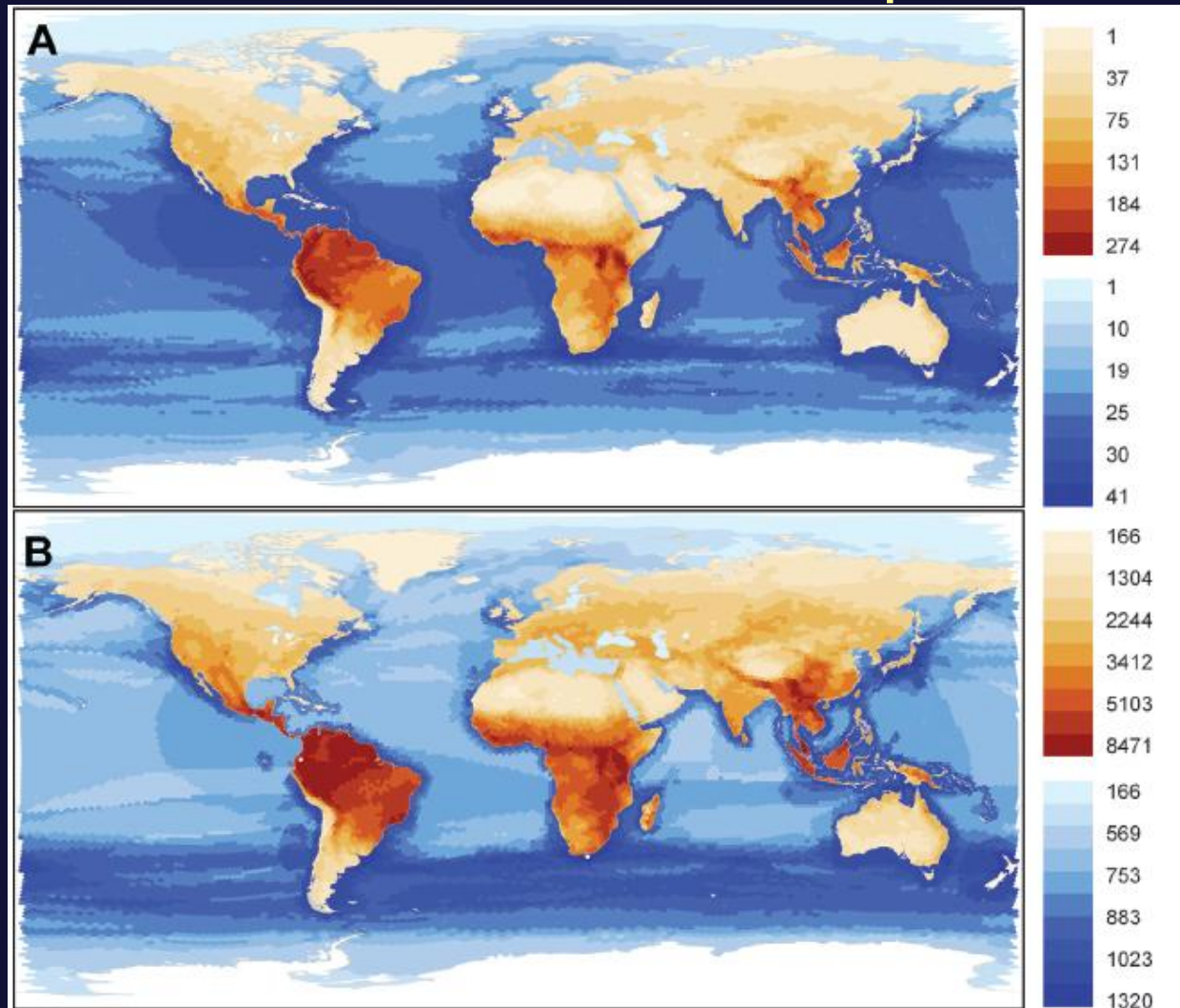
- *large* loss of PD or evolutionary potential



red = surviving evolutionary potential



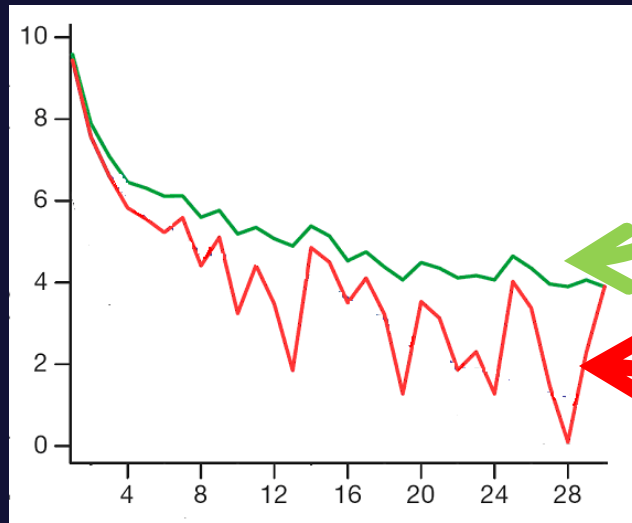
Total PD looks the same as total species diversity



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

but see Forest et al *Nature* (2007) – marginal gains/losses matter

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



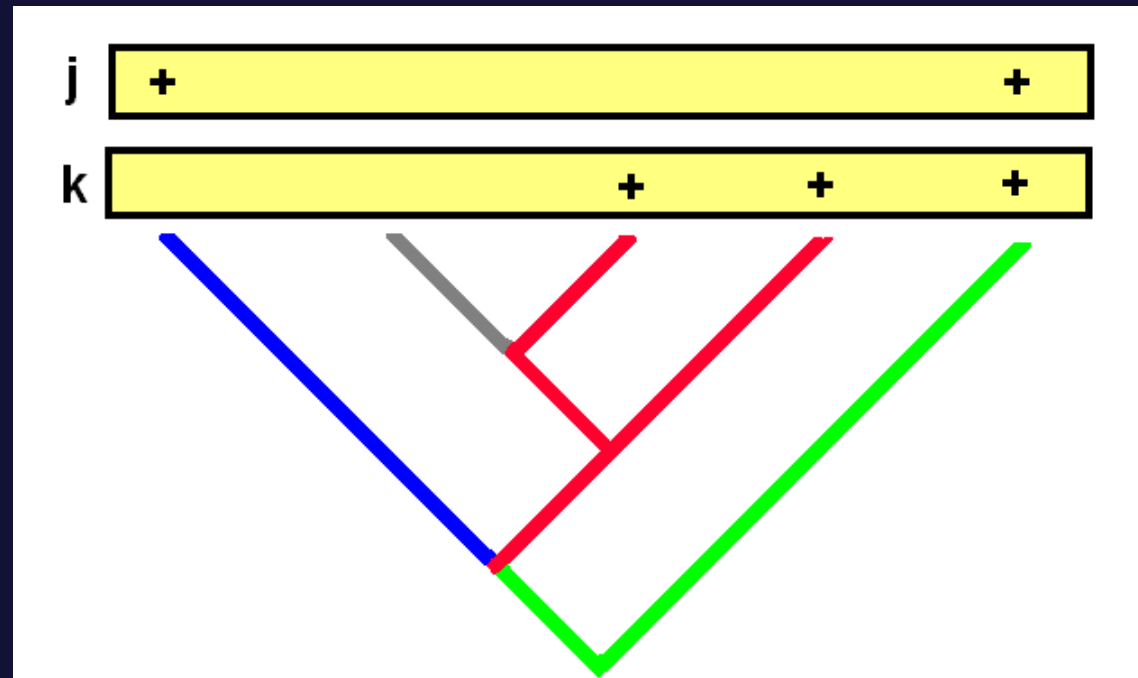
the PD that you could gain
does
not
match
the PD that you do gain



Forest et al
Nature 2007

PD complementarity shows how much branch length represented by an area was not already represented.
We also can compare two areas.

$$\text{PD Dissimilarity} = \frac{B + C}{2A + B + C}$$

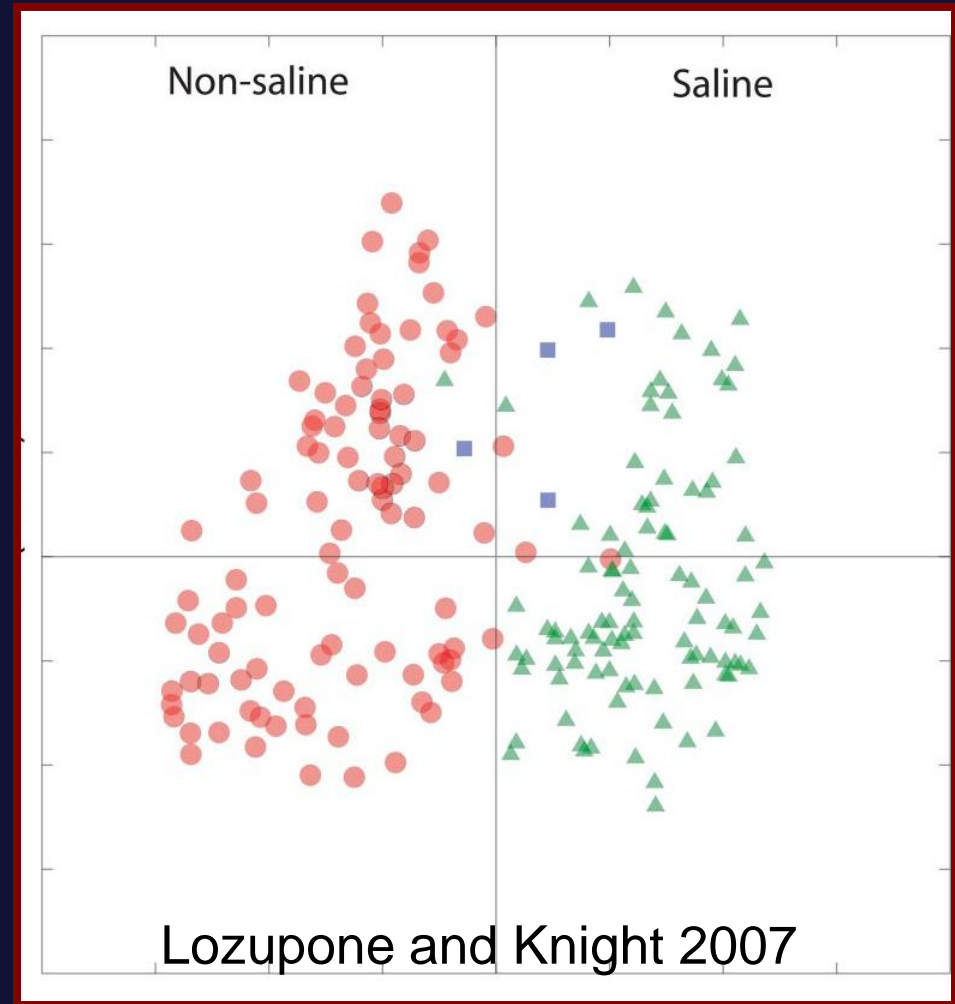


See
Lozupone et al 2005
Ferrier et al 2007

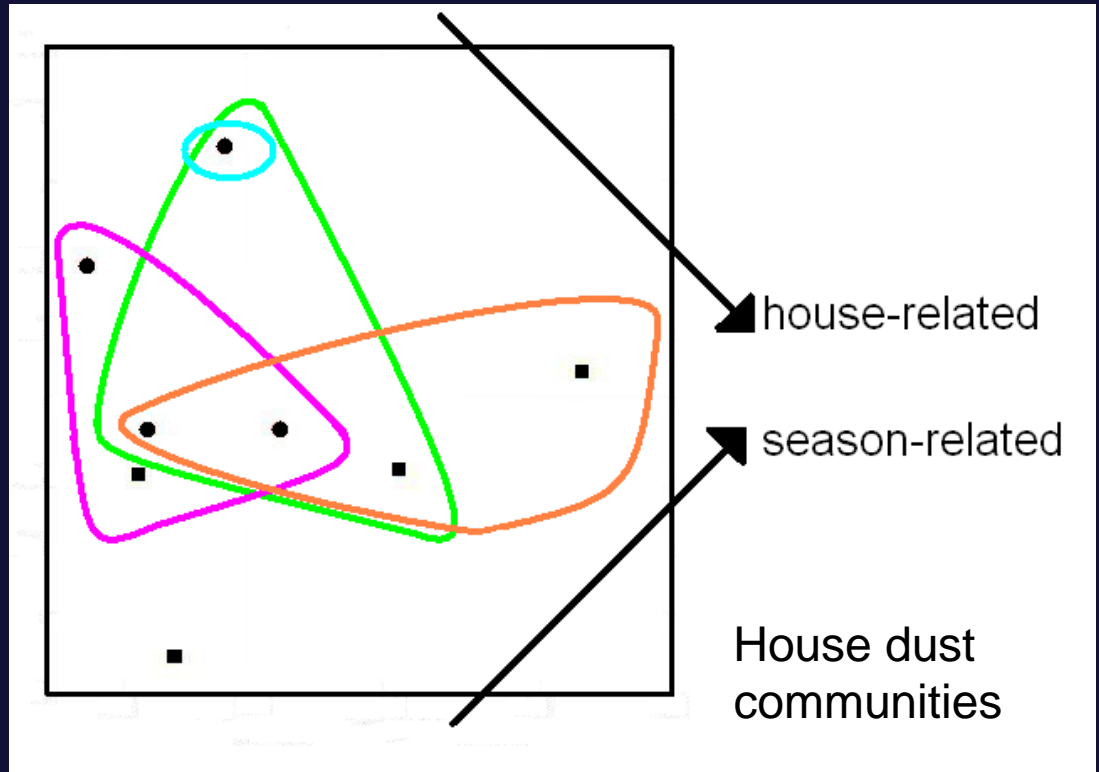
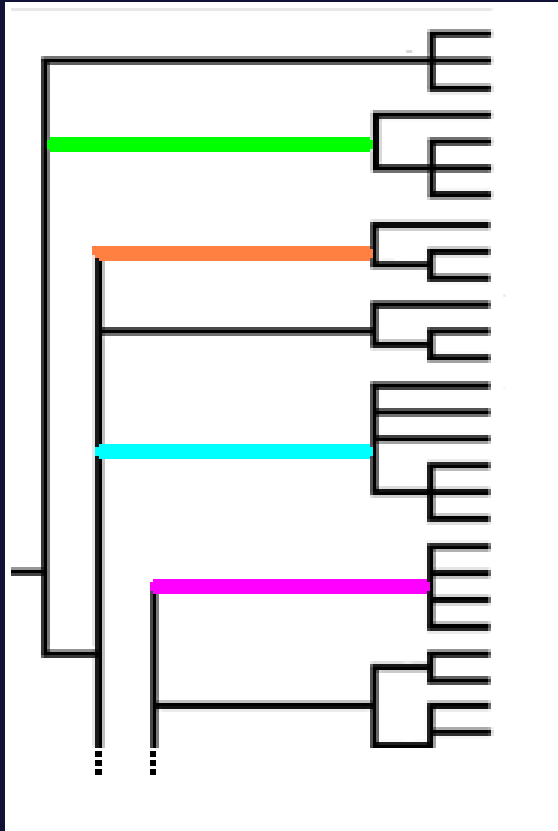
j and k dissimilar if there
is lots of red and blue

These dissimilarities can be calculated for all pairs of sample sites.

exciting examples include microbial ecology work using UniFrac (Lozupone and colleagues), where ordinations reveal key environmental factors/gradients e.g. global scale bacteria



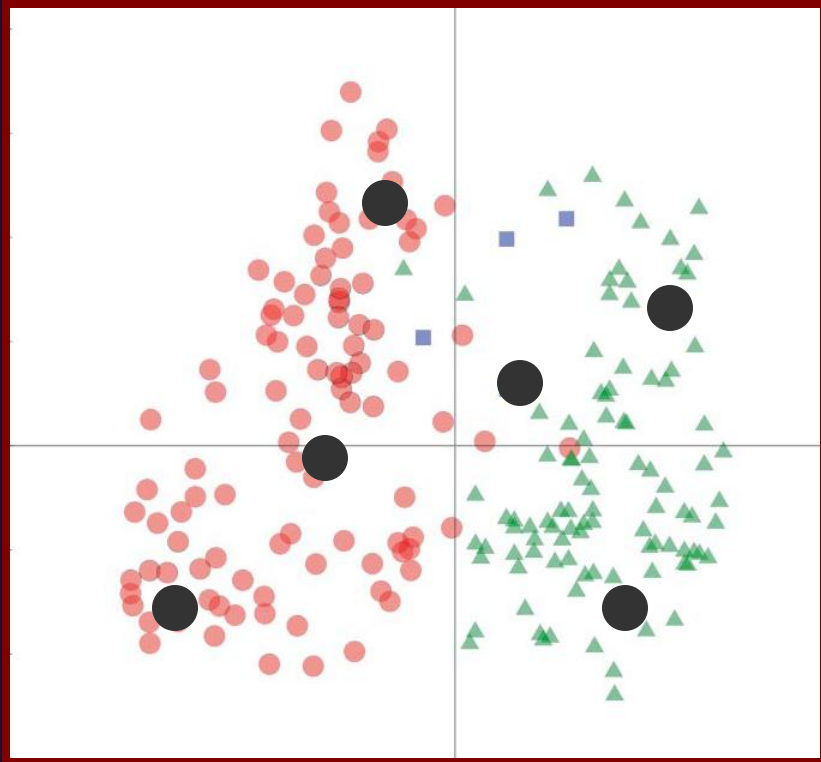
Phylogeny helps find important gradients, because even deeper branches have unimodal response to gradients



Faith, D. P., C. A. Lozupone, D. Nipperess, R. Knight 2009

The Cladistic Basis for the Phylogenetic Diversity (PD) Measure Links Evolutionary Features to Environmental Gradients and Supports Broad Applications of Microbial Ecology's "Phylogenetic Beta Diversity" Framework. *International Journal of Molecular Science*

The link from phylogenetic beta diversity to unimodal response of features/lineages means that ED method indicates diversity gains and losses





Using PD dissimilarities and ED, we have a “lens” for interpreting remotely sensed changes in land condition for GEO BON

Myobatrachid frogs (150 669 records, 0.01 degree sites) Classification based on predicted PD dissimilarities




Williams, Ferrier, Rosauer, Faith et al. Report for Department of the Environment, Water, Heritage and the Arts; for further information, contact Dan Rosauer

conservation priorities for species based on PD -




HomeAboutSpeciesConservationCommunityEDGE BlogSupport


Welcome to the EDGE



► TOP 100 AMPHIBIANS




Rank 2 Long-beaked echidna [View Species](#)





► TOP 100 MAMMALS

There's still time to save species on the EDGE




Welcome to the EDGE of Existence

Discover the world's most extraordinary threatened species - frogs that give birth through their skin and

 Print this page  Email page

EDGE Blogs



Saiga population assessment in western Mongolia
4th Feb 09
While we have all been enjoying the unusual amounts of snow in the UK, one of our EDGE Fellows, Buuvei, has been braving much more severe winter conditions t... [Read](#)

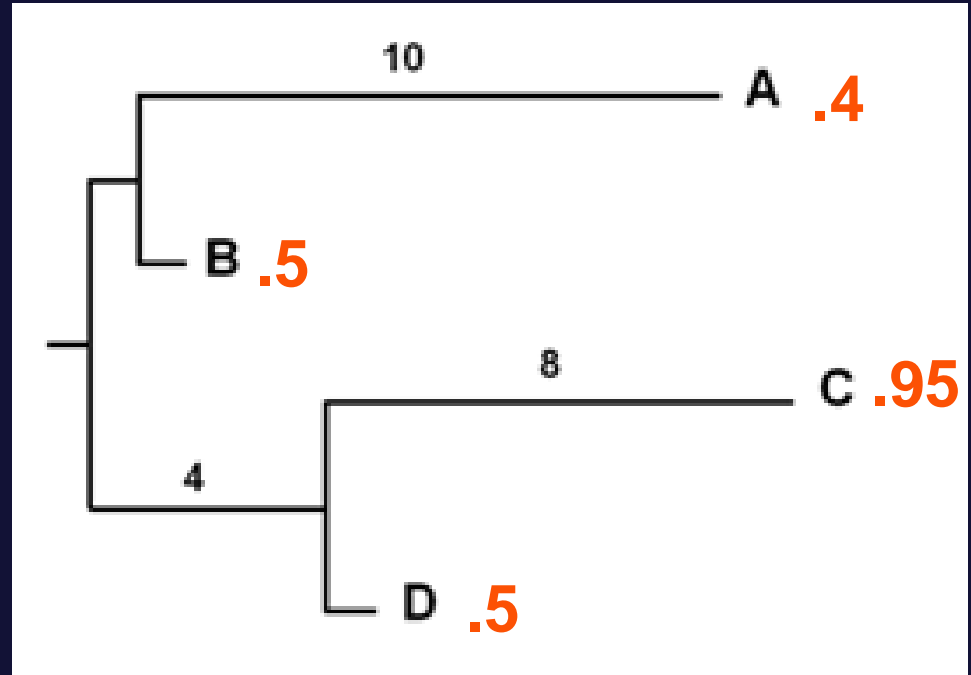
Zakhyn-Us Hay Crisis - update
23rd Jan 09
News just in from John Hare of the Wild Camel Protection Foundation that hay has been delivered to the captive breeding centre at Zakhyn-Us – just! John... [Read](#)

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

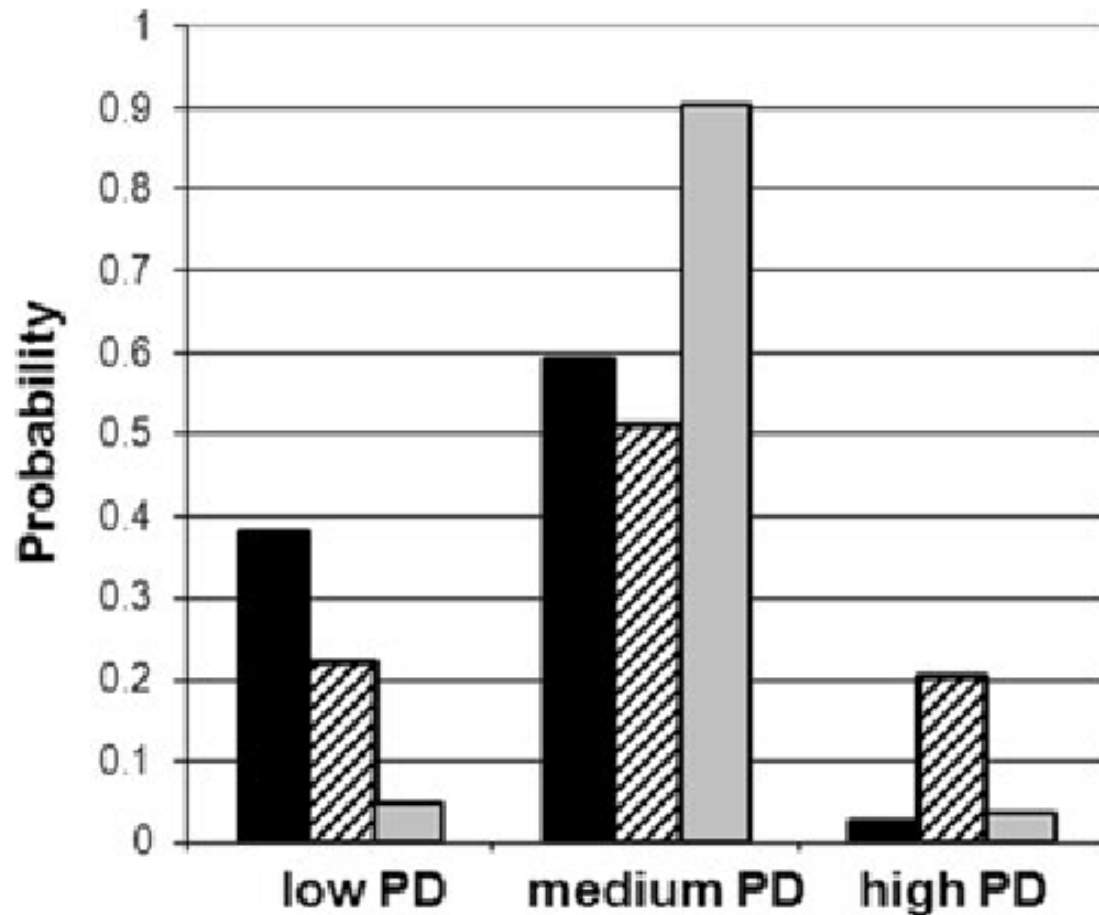
Faith DP (2008) Threatened species and the preservation of phylogenetic diversity (PD): assessments based on extinction probabilities and risk analysis. *Conservation Biology*

Phylogenetic risk analysis

Black =
current

Striped =
protect
species to
max
expected PD

Gray =
select
species to
avoid worst
case losses



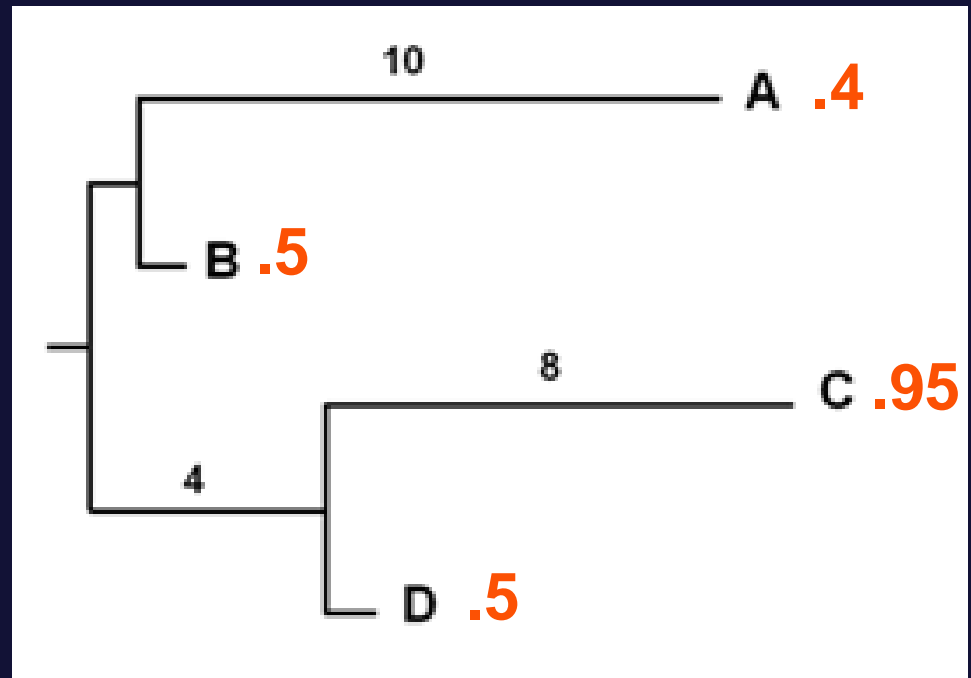
Faith DP (2008) *Conservation Biology*

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

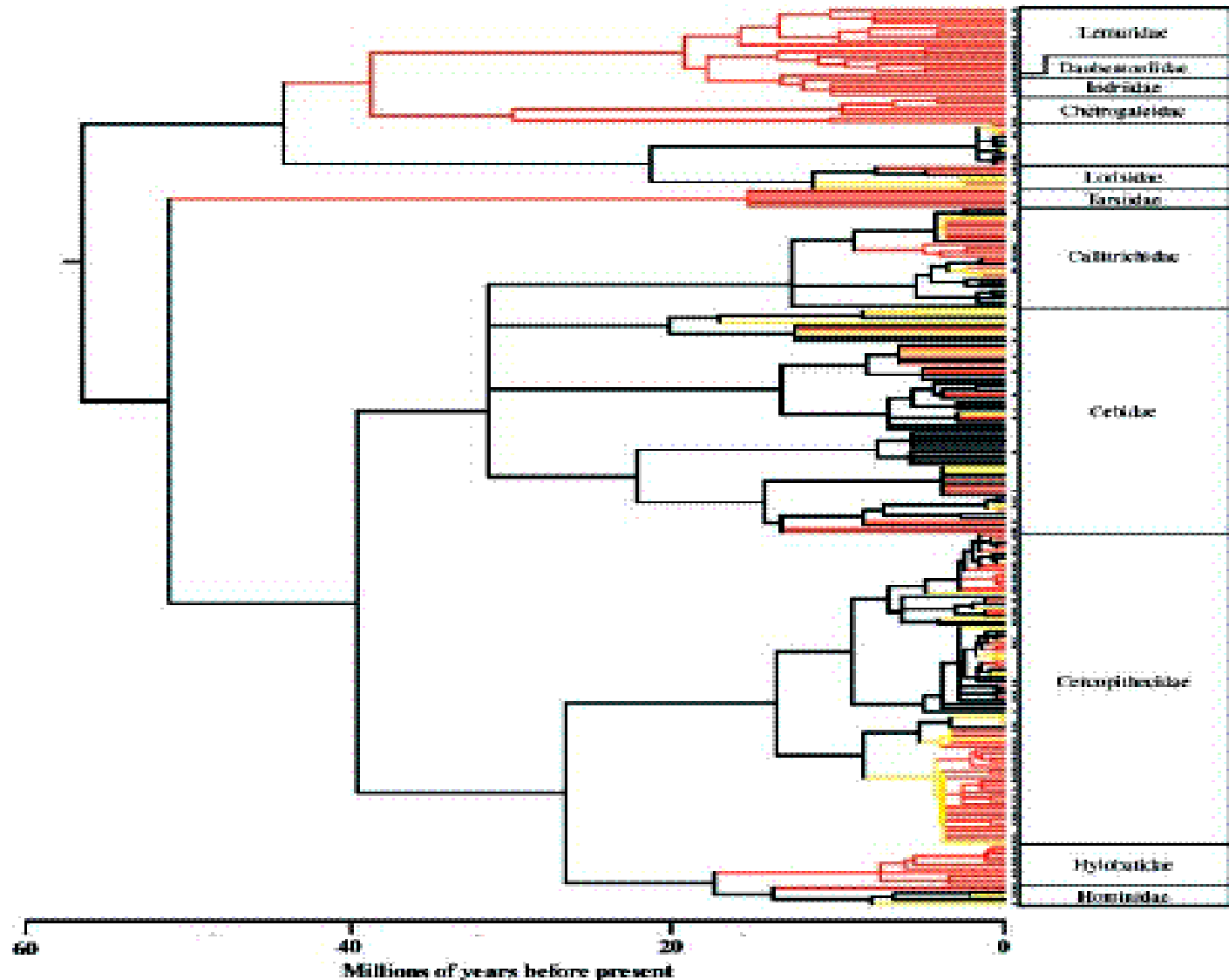


Can estimate “expected phylogenetic diversity”

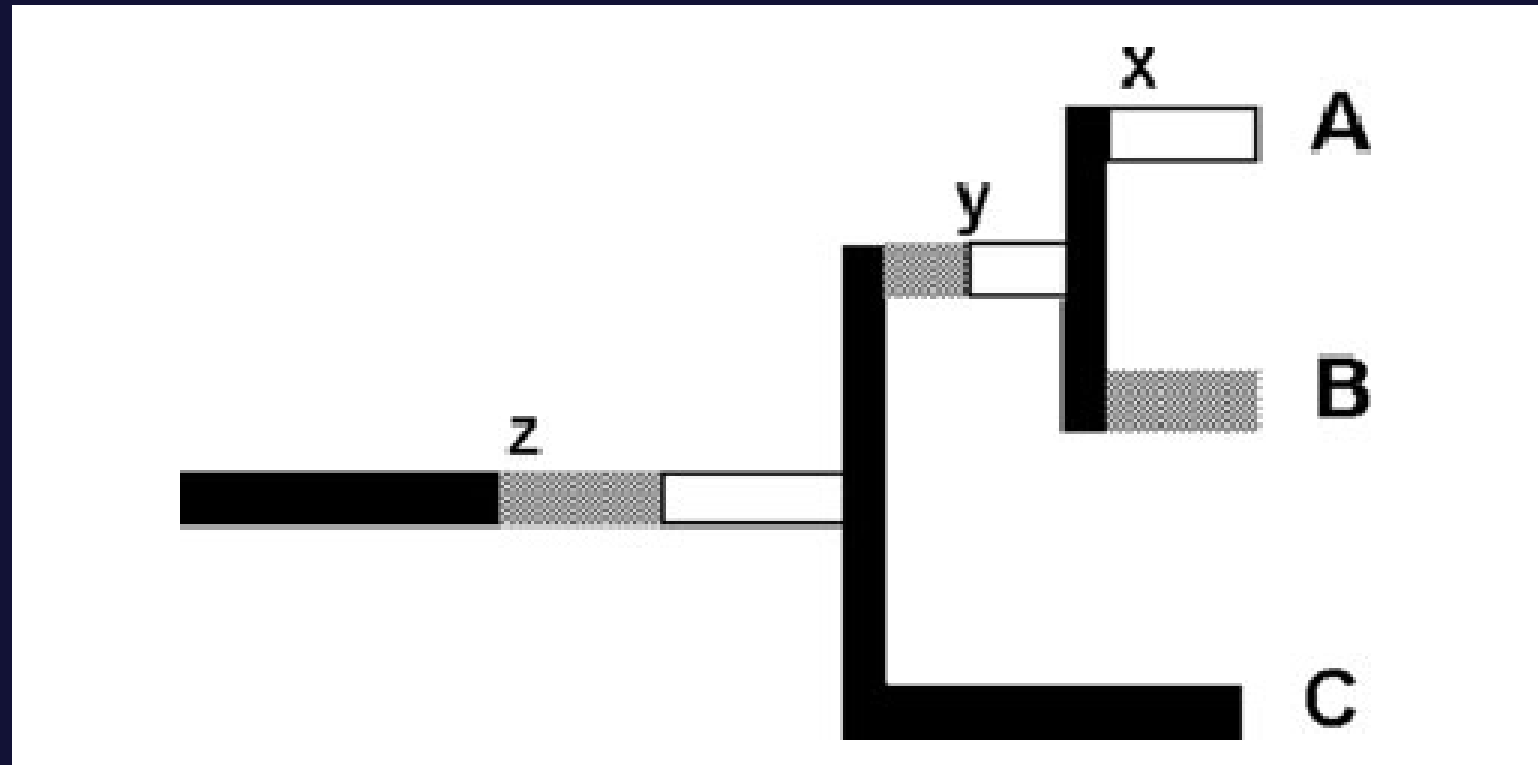
PD50 indices: for any species – what is the expected PD loss under extinction, assuming all other species have 50-50 chance of persistence? (see FISHBASE)

phylogenetic or PD-endemism

e.g. red branches restricted to hotspot regions



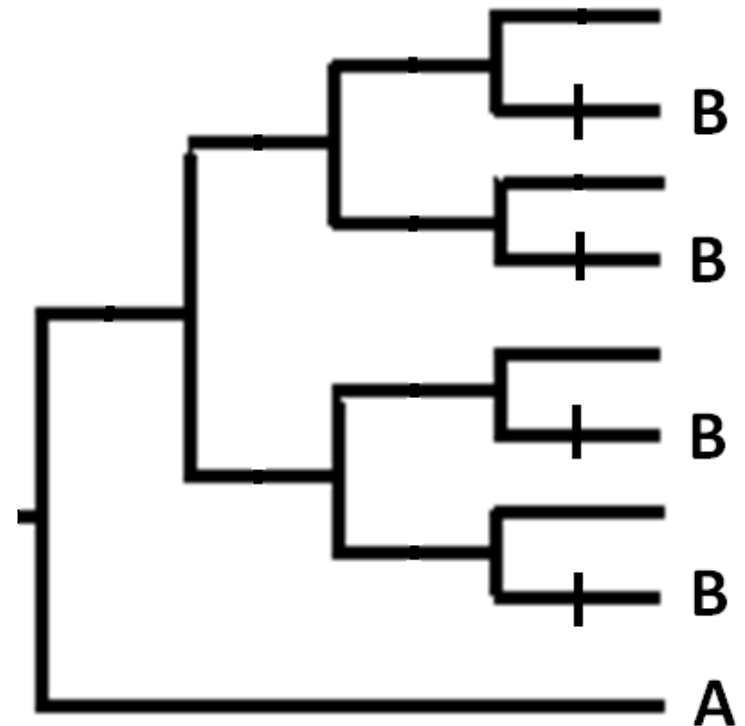
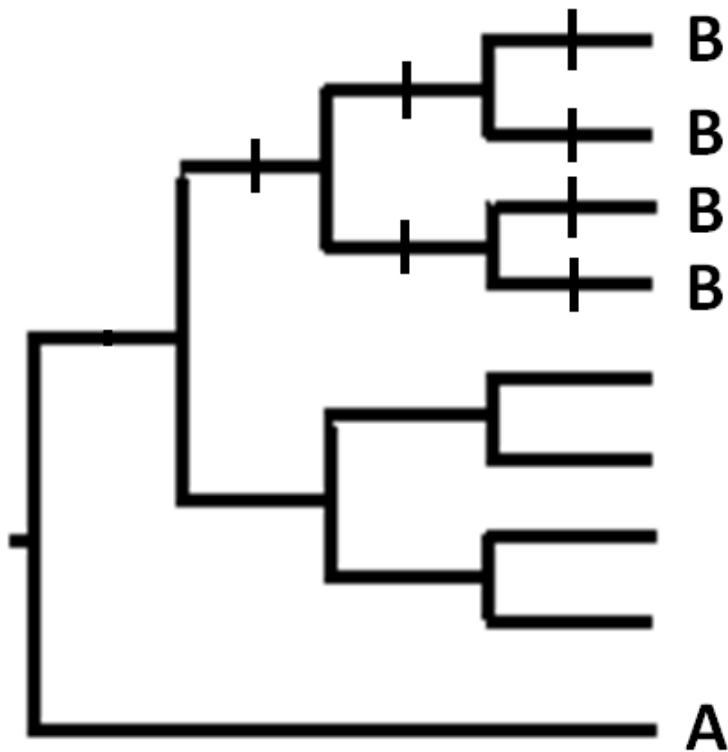
A weakness of EDGE calculations (and W-type indices for phylogenetic endemism measures)



W for a species is inversely related to the count of the number of groups on the phylogenetic tree for which the species is a member.

W scores may be divided-up among areas, and an area receives a score, equal to the sum of species' values

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



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,

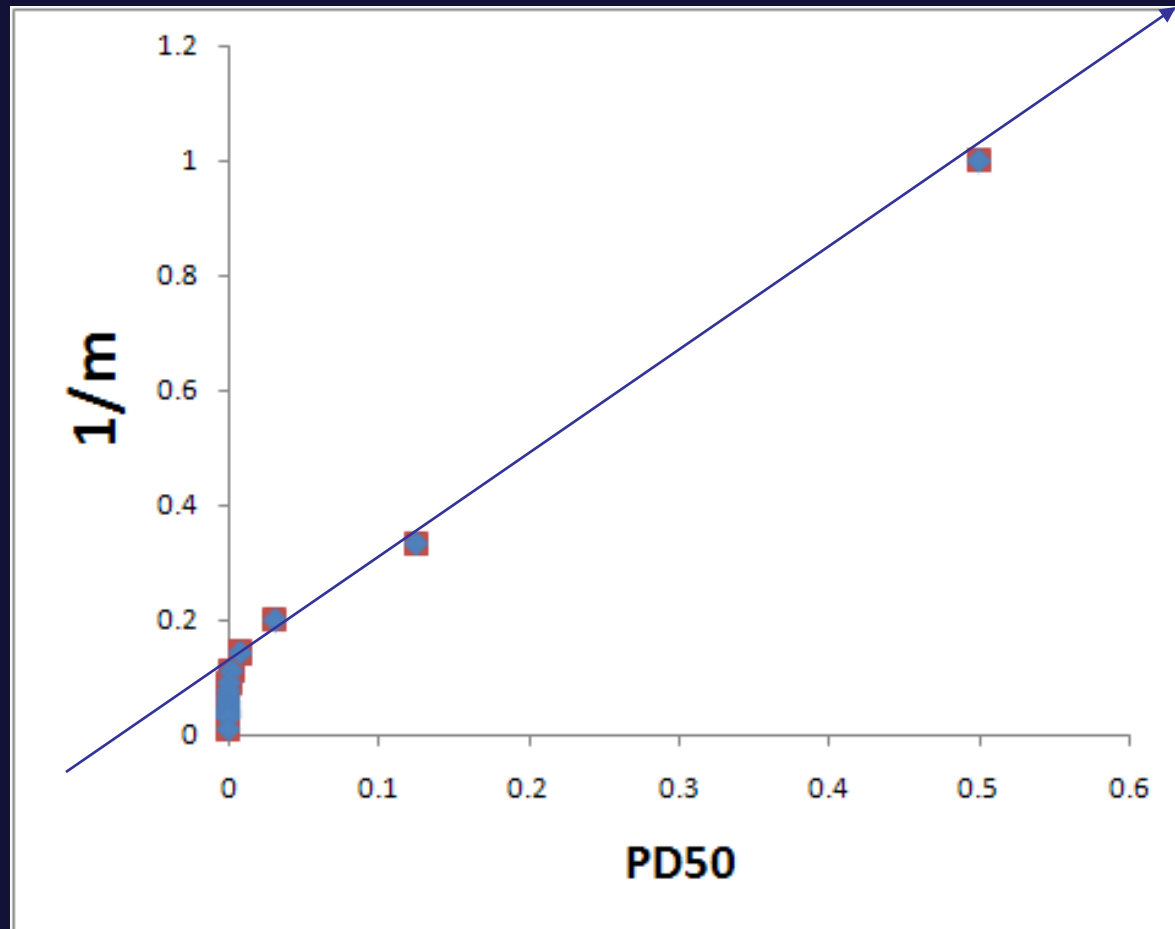
n_i = number of areas where branch descendants found,

$d_i = 1$ 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



Loss should be large to extent that area has long branches
 found in few descendants and few other areas
 - but this index does not integrate information about
 number of descendants of a branch

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

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

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:


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

consider q = probability of any species loss, ΔPDq =
expected PD before loss - expected PD after loss

$$\Delta PDq = \sum_i \left(Li \times q^{n_i} \times \left\{ \left[\prod_{j=1}^{k_i} q^{-1/m_i} \right] - 1 \right\} \right)$$

$\Delta PDq =$

expected PD before loss - expected PD after loss


$$\Delta PDq = \sum_i \left(Li \times q^{n_i} \times \left\{ \left[\prod_{j=1}^{k_i} q^{-1/m_i} \right] - 1 \right\} \right)$$


area has big score if:

it has **long branches** that have few descendants
(and those descendants are found in the area)

the branches are not found in many other areas

$\Delta PDq =$

expected PD before loss - expected PD after loss

$$\Delta PDq = \sum_i \left(Li \times q^{n_i} \times \left\{ \left[\prod_{j=1}^{k_i} q^{-1/m_i} \right] - 1 \right\} \right)$$



area has big score if:

it has long branches that have **few descendants**
(and those descendants are found in the area)

the branches are not found in many other areas

$\Delta PDq =$

expected PD before loss - expected PD after loss


$$\Delta PDq = \sum_i \left(Li \times q^{n_i} \times \left\{ \left[\prod_{j=1}^{k_i} q^{-1/m_i} \right] - 1 \right\} \right)$$


area has big score if:

it has long branches that have few descendants
(and those **descendants are found in the area**)

the branches are not found in many other areas

$\Delta PDq =$

expected PD before loss - expected PD after loss

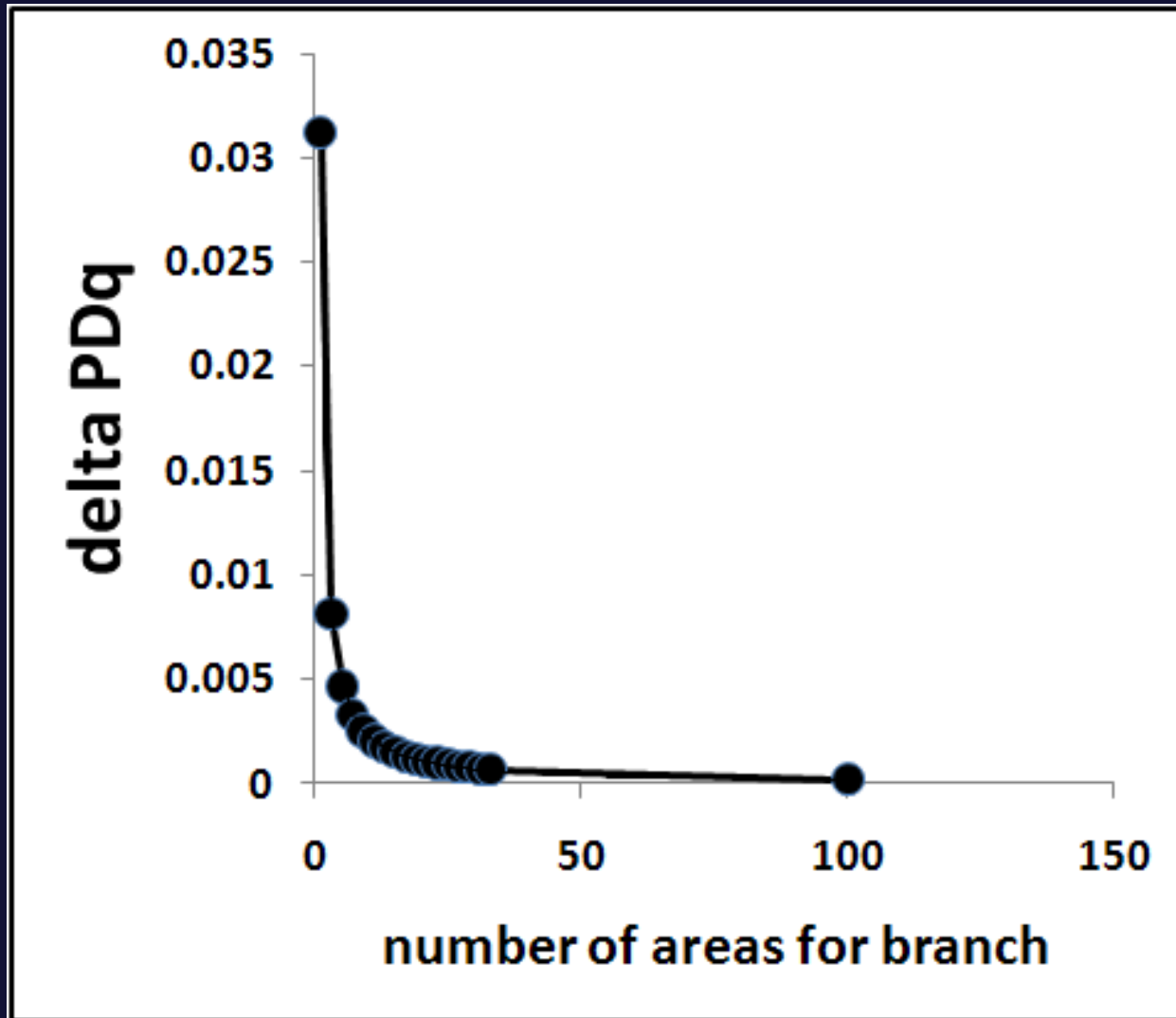
$$\Delta PDq = \sum_i \left(Li \times q^{n_i} \times \left\{ \left[\prod_{j=1}^{k_i} q^{-1/m_i} \right] - 1 \right\} \right)$$


area has big score if:

it has long branches that have few descendants
(and those descendants are found in the area)

the branches are **not found in many other areas**

Depending on q , the probabilistic PD endemism (ΔPD_q) index rapidly discounts branches found in many other areas



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

Faith et al.

Phylogenetic Diversity, Complementarity, and Endemism

25

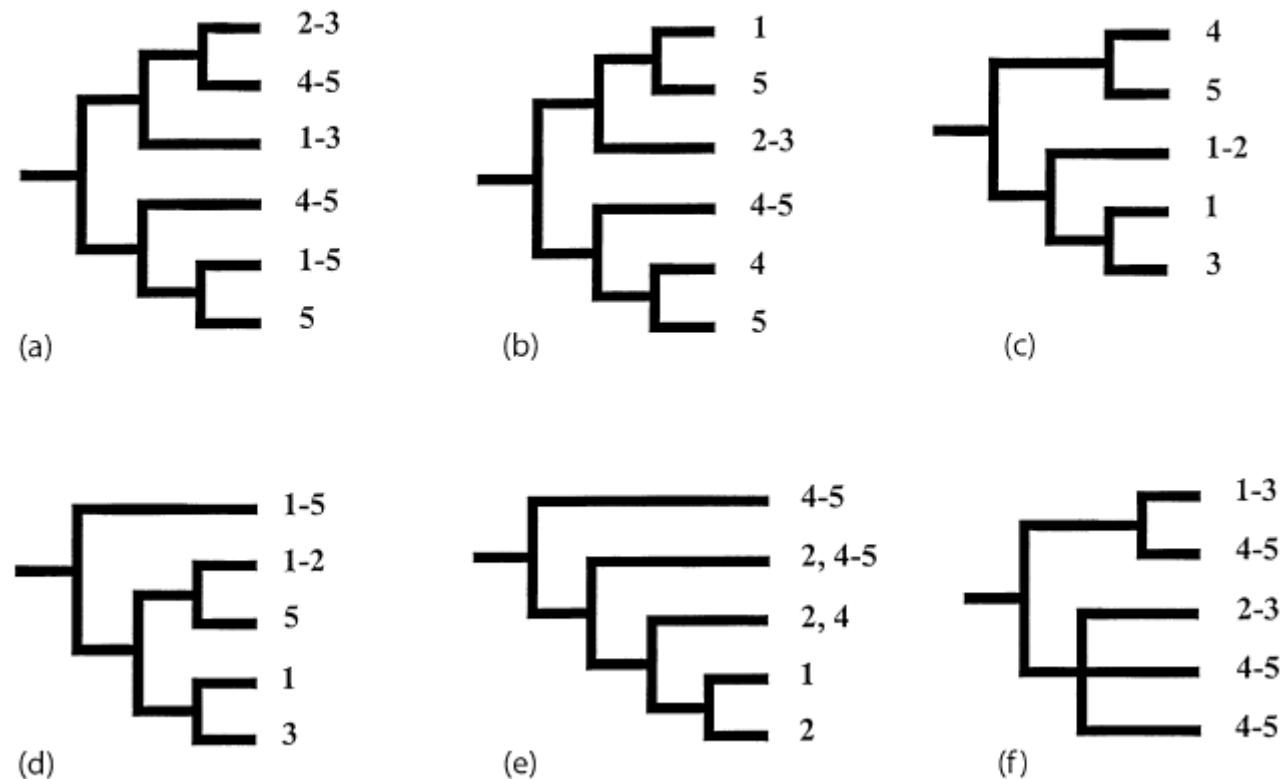


Figure 2. Estimated phylogenies for 10 taxonomic groups within Coleoptera for species in northeastern New South Wales, Australia. Branch lengths for each tree, as drawn, are arbitrary. For each phylogeny, species are labeled with numbers indicating areas of species distribution: 1, Barrington Tops massif;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

Failb et al.

Phylogenetic Diversity, Complementarity, and Endemism

25

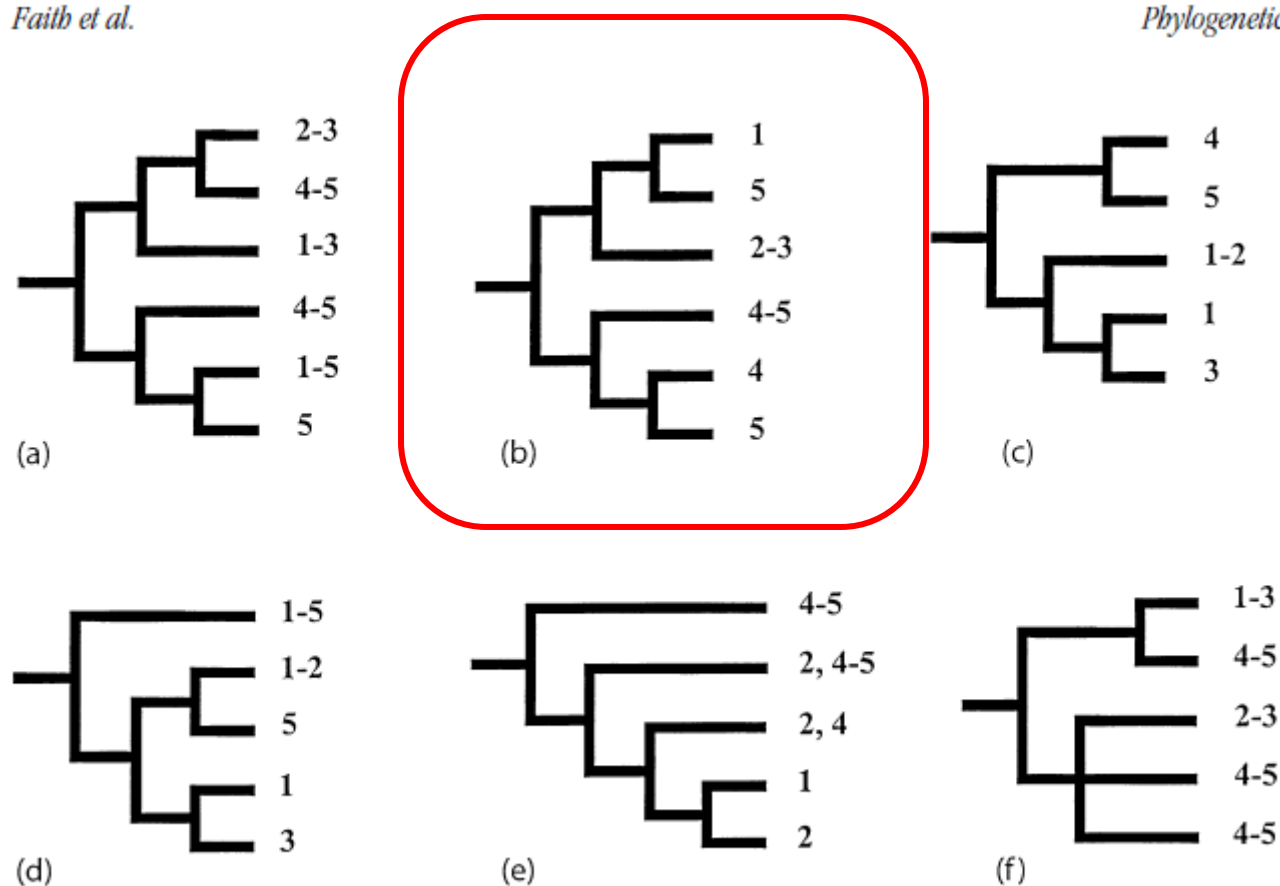


Figure 2. Estimated phylogenies for 10 taxonomic groups within Coleoptera for species in northeastern New South Wales, Australia. Branch lengths for each tree, as drawn, are arbitrary. For each phylogeny, species are labeled with numbers indicating areas of species distribution: 1, Barrington Tops massif;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

Some conclusions

- Phylogenetic ecology - take any conventional species-level index and re-express as a PD-based measure
 - e.g. The PD version of Bray Curtis dissimilarity has provided powerful method for interpreting microbial beta diversity
- Basic phylogenetic endemism useful , but sometimes want to give credit to areas with lineages with few descendants and found only in a small number of areas
 - ΔPD_q measures provide this
- PD measures integrating probabilities overcome weaknesses of methods using a static partitioning of phylogenetic tree among species
- GEO BON can help monitor phylogenetic diversity, both among species and among areas