# Phylogenetic diversity and endemism: metrics for identifying critical regions of conifer conservation in Australia

**Annasophie C. Lee, Brent Mishler**

*Department of Integrative Biology, University of California, Berkeley*

**Keywords: Australian endemics, biodiversity conservation, phylogeny, Biodiverse, ArcMap**

## Abstract

Accurately and sufficiently quantifying biodiversity is integral for conservation. Traditional metrics for measuring biodiversity, species richness (SR) and weighted endemism (WE), do not take into account the evolutionary history of organisms. Phylogenetic diversity (PD) addresses the shortcomings of SR by quantifying the evolutionary connections among the species present in an area. Phylogenetic endemism (PE) addresses the shortcomings of WE and represents the ranges of the branches of the evolutionary tree connecting the species in an area. Australia, with its advanced digitization of spatial reference data is the best model system for quantitative studies of biodiversity at present. I created a phylogeny for the 39 indigenous Australian conifer species using *matK* and *rbcL* sequences from *GenBank* and sequencing the 4 species for which there were no existing data. I used spatial data from Australia's Virtual Herbarium. More precise estimates of biodiversity can be used by conservation policy-makers.

## Introduction

Conserving global biodiversity, the variability between organisms, species, or ecosystems, is integral for conservation efforts (1,2). However, prioritization of critical species or regions for biodiversity conservation is a major challenge for conservation policy-makers from a number of perspectives. Historically, conservation efforts have often been focused on either conserving key species or regions (3). To identify key regions and species for conservation, measures of endemism have played a central role to quantify how restricted a species is to a given region. The degree to which species are restricted or widely dispersed is a strong predictor of extinction risk (4). Identifying these species at risk for extinction can be based on evolutionary history, geographic location, or a combination of the two. Geographically rare species are at greater risk of extinction (4), and phylogenetically rare species (5) contain disparate genetic information and contribute heavily to biodiversity, thus it is critical to examine the intersection of these subjects.

The quantification of biodiversity has historically been problematic and current metrics are problematic because they do not include an evolutionary perspective. For example, enumeration of species is hindered by the lack of a universal

agreed-upon species concept across researchers reflecting an arbitrarily decided level of genetic and morphological variation, which leads to inconsistency in taxonomic ranking or hierarchy (6,7). Additionally, inconsistencies in identification and discovery of species lead to false classifications that both under and overestimate biodiversity. More importantly, these issues with naming and identifying species are compounded when traditional biodiversity metrics are calculated without considering evolutionarily relatedness between species, and their dispersal from their geographic origins. Species richness (SR), the absolute number of species in a region, was developed to quantify the number of species in a region and weighted endemism (WE) quantified their level of endemism (8,9). However, SR and WE as measures of biodiversity consider only the terminal taxa of a phylogenetic tree, without considering the evolutionary relationships among them (10). Species vary in their evolutionary isolation and genetic diversity, and these differences give insight into how species may have evolved and which are most important for biodiversity conservation (11). SR and WE do not include information about how closely related species are, excluding relationships between sister groups given by the phylogeny. Consequently, these metrics are limited in their ability to describe biodiversity patterns as they are a more surface-level analysis of biodiversity as compared to one that incorporates the evolutionary perspective (10, 11, 12).

Diversity measures based on phylogeny, or the evolutionary relationships between species, have since been developed to address the shortcomings of descriptors such as species richness and species endemism. Phylogenies are derived from shared, homologous characters, or characteristics shared by all the descendants of a common ancestor, and are an indication of recently shared ancestry. Phylogenetic diversity (PD) and phylogenetic endemism (PE) are metrics that provide a more comprehensive view of diversity within and between species (10, 12, 13, 14). PD calculates the shared evolutionary history of specified taxa (10,14) and is largely resistant to taxonomic uncertainty, or the discrepancies in the identification of species, because it relies on robust hypothesis of evolution, derived from the shared homologous characters between species (15). PD has been utilized to understand global patterns of biodiversity, and is especially useful when the taxonomy of a clade is poorly understood (13). PE is a measure of the amount of shared evolutionary history between a set of branches on a phylogenetic tree in relation to how widespread the branches are geographically (10). WE is the sum of the inverse of the species' range found over a fixed area (9). PE, unlike WE, incorporates the ranges of all the branches of the tree connecting the species, not just the terminal branches (10). This weighted phylogenetic endemism provides a more comprehensive measure of the distribution of rarity than weighted endemism of species alone. PD and PE are more robust to changes in taxonomic classification than SR and WE, and PE analyzes endemism across a consistent spatial scale, regardless of previously defined geographic boundaries (10). These metrics provide evolutionary and genetic information necessary for making informed conservation-policy.

Calculating PD and PE requires a high resolution of spatial distribution information along with a highly resolved phylogeny. Australia is the best model system, at present, for this type of study due to the advanced state of digitization of herbaria voucher specimens and spatial reference data (16, 17). Australia's Virtual Herbarium (AVH), contains millions of records of spatial flora collections from Australia's major Herbaria. Additionally, Australia is important for global biodiversity conservation as it is rich with endemic species, resulting from its geographic isolation (9, 18). Conifers are also largely confined to either the Northern or Southern hemisphere, specifically extant species of Araucariaceae, Podocarpaceae, and the Callitroideae (the sister group to Cupressoideae), fossil records also indicate that these trends have persisted throughout time (19). Although metrics such as species richness and species endemism have been calculated for many conifer species in Australia (1) calculation of diversity metrics from an evolutionary perspective using PD and PE remains to be accomplished.

The main objective of this study was to calculate and visually display diversity metrics that couple phylogenetic and spatial information. I calculated PD and PE to identify regions of Australia most densely populated with phylogenetically rare conifers and compared these results with Australian natural reserves to identify regions of phylogenetic rarity that are not currently being protected. I hypothesized that species which are evolutionarily distant will be more geographically distant and closely-related species will be spatially clustered, because of habitat requirements (20). Additionally, I evaluated the relationships between PD, PE and traditional diversity metrics such as SE and WE. I expected PE to be correlated with WE; however I expected WE to fail at consistently predicting areas of high PE (10). These results will prove valuable to informing conservation-policy makers regarding critical regions of conifer conservation.

## Methods

### *Spatial data acquisition*

I studied the 39 indigenous species of conifers in Australia (Appendix A, List 1). To obtain specimen locations, I used data from AVH (http://avh.ala.org.au/). The AVH is a digital database containing 75% of the 6 million specimens of plants, algae and fungi that have been collected by Herbaria in Australia. I downloaded a total of 12,300 Australian endemic conifer species datapoints and then used Google Refine version 2.5 (http://code.google.com/p/google-refine/), to clean the dataset and remove non-conifer records, foreign collections (as well as Norfolk and Macquarie Islands), and any naturalized specimens grown in a botanic garden or otherwise. I reconciled the taxonomy against a classification for extant conifers with the Australian Plant Census (APC)

**Figure 1. Spatial location of individual conifer specimens.** Specimens collected using AVH database.

(http://www.anbg.gov.au/chah/apc/index.html) and corrected any misspellings. I trimmed records without geographic coordinates from the dataset. I then transformed the latitude and longitude values of the remaining records into xy meter coordinates using the Albers projection which corrects for inconsistencies in grid size of latitude and longitude near the earth's poles. This cleaned dataset contained 7300 spatial records (Fig 1)

### *Molecular data acquisition: Phylogeny*

I used two genes, *matK* and *rbcL* to create a phylogeny, using both existing and new sequence data. *RbcL* is commonly referred to as the "universal barcode" for plants; however using two genes, both *matK* and *rbcL*, is more informative and created a more complete and accurate phylogeny (Quinn et al. 2002). I searched the online database *GenBank* (http://www.ncbi.nlm.nih.gov/genbank/) using scientific names of each of the 40 species in my study (39 indigenous species and one outgroup, *Gingko biloba*). I noted which sequences were unavailable in *GenBank* and saved the accession numbers of the available sequences (Appendix A). Once I identified which species were missing, I collected plant tissue for *Callitris baileyi, Callitris monticola, Callitris oblonga, Callitris columellaris, Actinostrobus acuminatus* and *Microstrobos niphophilus* at the Royal Botanic Garden, Sydney. I located the species I needed in the botanic garden and cut a piece of fresh leaf from which to extract DNA.

After I cataloged the plant tissue in the Royal Botanic Garden Herbarium's collection, I prepped my tissue samples for DNA extraction by sealing them in a silica gel filled box to desiccate them. I then performed DNA extractions using a Qiagen DNEasy Kit (Germany,www.qiagen.com) with minor modifications. These modifications were: using 1 zirconia bead and 5 mg sand instead of 50 $\mu$L small zirconia beads, not using any liquid nitrogen, using the lyser (written "bead-beater" in kit) for 25 seconds, incubating at 65°C for 40 minutes, and incubating the products of buffer AE and DNA for 10 minutes .

Once I extracted DNA from the leaf tissue, I amplified the regions *matK* and *rbcL* using PCR. I performed the standard procedure using the primers Forward TX2 and Reverse TX4 to amplify *matK* regions and the primers Forward *rbcL*_1 and Reverse *rbcL*_635. I ran a program called Immolase 50°C on the Thermocycler (Corbett Life Science, Palm-Cycler) for 2.5 hours. Then I loaded the product into wells on gels and ran electrophoresis on the gel with indicator and gel red at 300 W for approximately 10 minutes, checking to see the movement of the bands periodically. I then transferred the plates to a UV hood and visualized the plates. After taking note of which trials were successful, I collected the PCR products for sequencing. I then sent the PCR products to the Genetic Sequencing Lab on the UC Berkeley campus. The sequenced products were then sent back to me as a data file.

### *Phylogeny construction*

To create the Australian conifer phylogeny, I acquired DNA sequences from the processes outlined above and used the default settings for the MUSCLE alignment in Geneious (http://www.geneious.com/) to align the sequences for each gene region, *matK* and *rbcL*. Once I aligned the genetic sequences, I deleted any unreliable end pieces that were unlikely to represent *rbcL* or *matK* gene regions. I chose one *matK* and one *rbcL* sequence to represent each species using the following criteria,

known as taxon priming: longest sequence, a sequence that withstands a cluster analysis, and Australian in origin. I then created a concatenated matrix including *rbcL* and *matK*, a total of 2783 base pairs, and used the default parameters in GARLI (Genetic Algorithm for Rapid Likelihood Inference) version 0.951 (https://code.google.com/p/garli/) to create a Maximum Likelihood phylogeny (Fig 2). I then compared the relationships in the phylogeny I created with previously published conifer phylogenies (e.g., 19).

### *Biodiverse: Spatial Location and Phylogeny*

*Biodiverse v 0.17* (http://code.google.com/p/biodiverse/) is a program that uses a phylogeny and specimen level spatial data to create a map of the occurrence of species across a region and calculates SR, WE and phylogenetic metrics PD and PE. SR and WE require only spatial data, whereas PD and PE require spatial and phylogenetic data. I loaded the cleaned spatial data I acquired from AVH into *Biodiverse* which displayed a map each species' occurrence (Fig. 1) and the phylogeny I created from the gene regions *matK* and *rbcL* .

First, I calculated species richness—defined as the number of species in an area (here represented by 50,000 $m^2$ grids). Second, I calculated PD (Eq 1, 21), which is calculated by summing the branch lengths on the phylogenetic subtree connected the species in a particular gird. Third I calculated PE, defined as PD weighted by the inverse of the branchlength's ranges. PE incorporates the spatial range of the phylogenetic branch lengths down to the root of the phylogeny (10). For example, if a widely distributed taxon is sister to a narrowly distributed (highly endemic) species, the highly endemic species will be negatively weighted by its sister and the PE score of the pair will be lowered.

$$\mathrm{PD} = \sum_{c \in C} L_c \qquad \text{(Eq. 1)}$$

where Lc is the length of branch c and C is the set of branches in the minimum spanning path connecting the species (Rosauer et al. 2009).

$$\mathbf{PE} = \sum_{\{c \in C\}} \frac{L_c}{R_c} \qquad \text{(Eq. 2)}$$

Where variables are defined as above, and Rc is clade range, the combined ranges of the descendant taxa of branch c, so that overlapping areas are considered

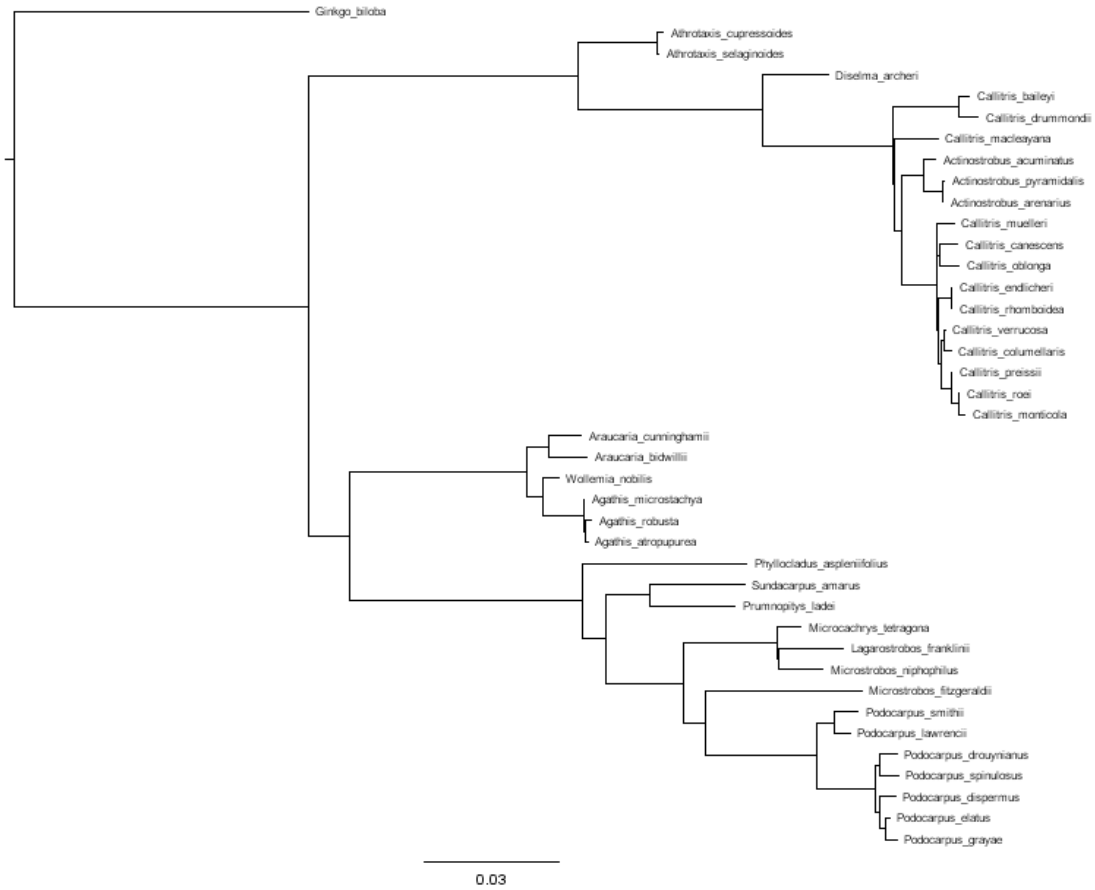



**Figure 2. Maximum likelihood phylogeny of endemic Australian conifers.** Derived from *matK* and *rbcL* gene regions and calculated using *GARLI v. 0.951.*

only once (Rosauer et al. 2009).

To discern any correlations between SR and PD, I created a scatterplot of SR as a percent of total number of species against PD. I performed the same calculation for WE versus PE, to graphically display any correlation between the two metrics. I also calculated the correlation coefficient for each relationship ($r^2$).

### *Spatial analysis*

To determine whether areas of significant PD and PE were correlated with protected regions in Australia, I overlaid map layers of Natural Parks and Reserves in Australia using *ArcMap v 10.1* (GISESRI). I gathered the data layers from the Atlas of living Australia (http://spatial.ala.org.au/) and loaded the data layers into *ArcMap*, projected them, if they were not already in the projection GDA94 / Australian Albers. I then projected the *Biodiverse*-exported ASCII grid files into the same projection to visualize properly. I clipped the data if it contained more data points than the continent of Australia. Then I symbolized the data to display the Australian Protected regions shapefile (CAPAD 2010) and overlaid an outline of the shape of Australia in GDA94/Australian Albers Projection to display the continent's bounds. I used this visualization process to discern any correlations or patterns in the data layers over the randomization maps of RPD, RPE and super-endemism.

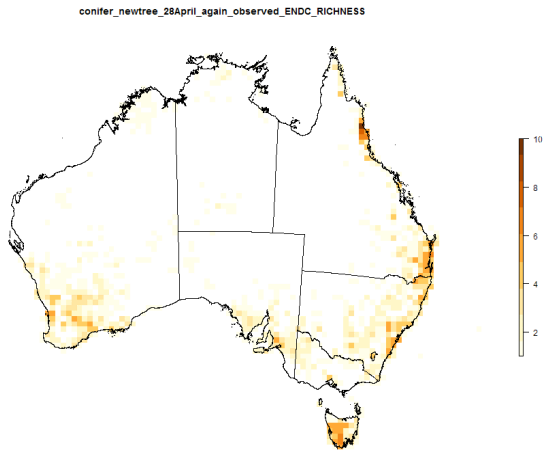
## Results

### *Study organisms and study site*

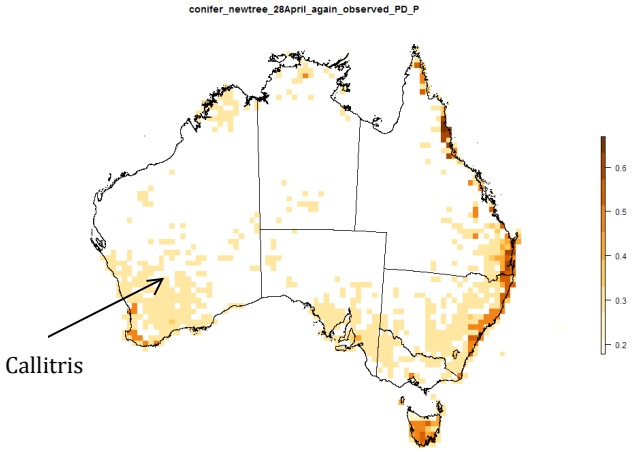
The phylogeny I created is fully resolved, and provides a robust hypothesis of the evolutionary relationships between Australian endemic clades. However, it probably includes an incorrect relationship: *Microstrobos niphophilus* probably belongs in the same clade as *Microstrobos fitzgeraldii* (19). For the purposes of these calculations it does not make a difference, because both PD and PE take into account branch lengths, and the erroneous branch is very short. Fig. 2 is the result of a maximum likelihood phylogenetic tree for the 39 conifer species, rooted on the outgroup, *Gingko biloba*.

### *Biodiverse: Geographic Location and Phylogeny*

I found that species richness was highest in Tasmania and on the Northeast coast of Australia (Fig 3a). PD was more scattered than SR, but also clumped in Tasmania and on the East Coast (Fig 3b). SR was fairly strongly correlated with PD ($r^2$=0.75), where r




**B**




**C**

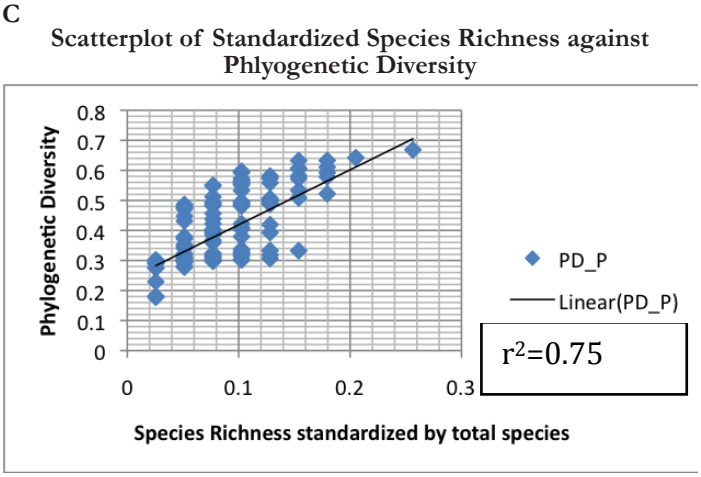
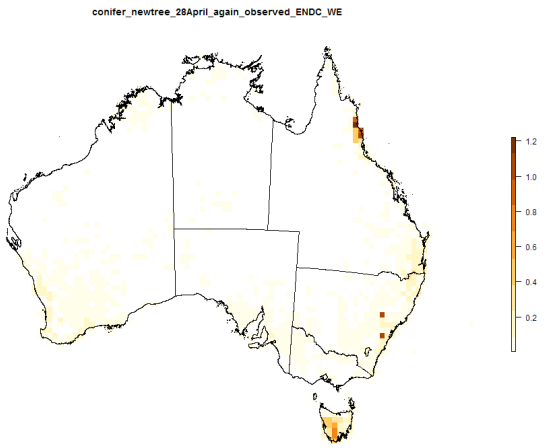
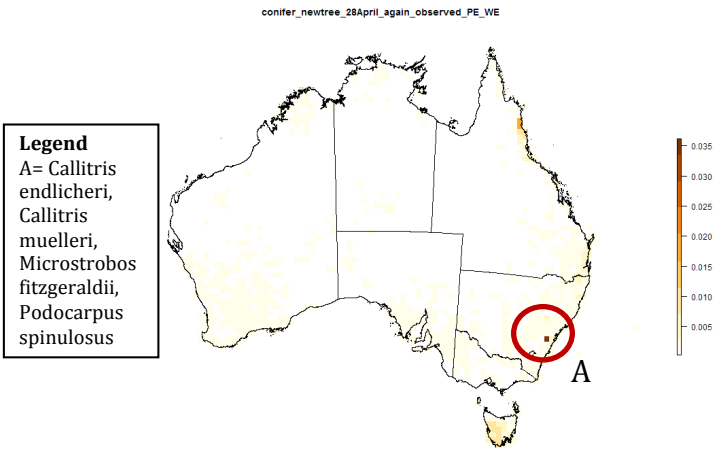



**Figure 3. (A) Species Richness** Species richness of endemic conifers species in Australia. Red regions represent high species diversity.regions and calculated using GARLI v. 0.951. **(B) PD of conifers** Phylogenetic Diversity of conifer species in Australia. The dark red regions, primarily on the East Coast and Tasmania, represent high levels of PD. The genus *Callitris* was widely distributed, especially on the West coast of Australia. **(C) Species richness (%) against phylogenetic diversity weighted by branch lengths, with a best fit line.** Note that data were not normally distributed, so $r^2$ is a residual value, thus I did not include a significance value.

A




B




C

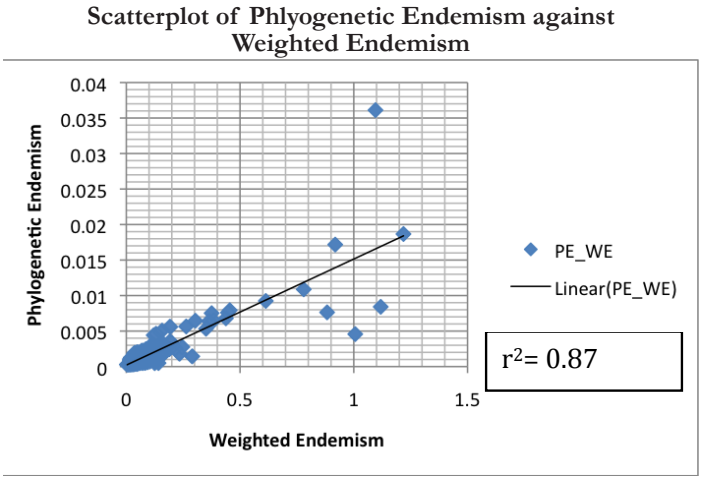



**Figure 4. (A) Weighted endemism of conifers** Weighted endemism of endemic Australian conifers. **(B) Phylogenetic Endemism of Australian conifers.** Dark regions represent high PE (PE>0.035). The grid cell labeled A contains the genera *Callitris, Microstrobos* and *Podocarpus*. PE is also relatively high on the Northeast coast. **(C) Weighted endemism against phylogenetic endemism with a best fit line.** PE is overall strongly correlated with WE, but this correlation does not hold for some values of high PE or high WE. Note that data were not normally distributed, so $r^2$ is a residual value, thus I did not include a significance value.

represents the residual value, as the data were not normally distributed. Tasmania has an especially high PD score and contains species that are distantly-related, *Athrotaxis, Diselma, Lagarostobos, Microstrobos, Phyllocladus* and *Podocarpus.*

WE was concentrated primarily in Tasmania and along the Northeast coast (Fig 4a). PE was not as high in Tasmania, but also was concentrated along the Northeast (Fig 4b). WE was highly correlated with PE, but underestimated some regions of high PE. For example, the grid cell which contained the highest PE value, 0.0361, was underestimated by WE (Fig 4c). This grid cell contained *Callitris, Microstrobos,* and *Podocarpus*, which are not sister terminal taxa on the phylogeny.

### *ArcGIS Analysis*

After calculating PE, PD with randomizations and visualizing with CAPAD 2010 Protected Regions, I found that the majority of regions with significantly high (p value>0.95) PD were protected (Fig 5). Those cells that contained significantly high PD values are concentrated in areas that are heavily protected, such as the far Northeast and Tasmania.

## Discussion

Accurately and sufficiently quantifying biodiversity is essential for conservation efforts. In this study, I explored biodiversity metrics that quantified the spatial distribution of evolutionary history of Australian endemic conifer species in comparison to traditional metrics which do not take evolutionary history into account. SR and PD were largely correlated, with some exceptions where SR did not predict PD values accurately. WE and PE were also largely correlated, but that correlation broke down for some high values of WE or PE. The spatial and phylogenetic analysis yielded that most regions, high or low with PD and PE, are currently being protected as reserves under Australian law.

### *Phylogenetic Metric Performance*

*Regional trends in species richness, endemism vs. PD and PE*

As a whole, the continent of Australia had relatively low PD values compared to a random distribution, which could be due to biogeographic barriers to dispersal and diversification (12). SR and PD were

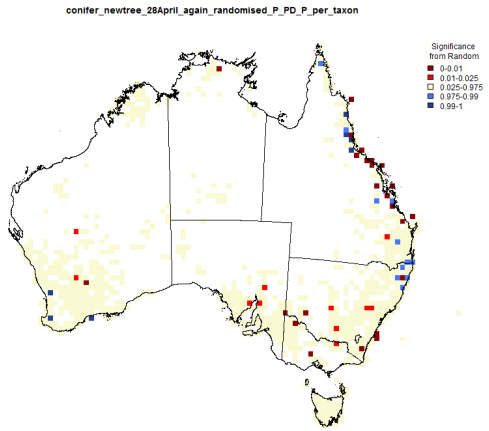



**Figure 5. Randomization of PD** Significantly low PD (red) is scattered throughout the south and on the Northeast coast. There are fewer regions of significantly high PD (blue), and they are concentrated primarily on the Northeast coast.

largely correlated, which one would expect (20), given that the more terminal taxa that are sampled from a specific grid cell, the more of the phylogenetic tree is sampled. However, some regions had more or less PD than predicted by their SR (Fig 3c). This correlation was weak for intermediate levels of species richness and PD (Fig 3c). In most cases, SR underpredicted PD, meaning that there were more distantly –related taxa in that grid-cell than expected given SR count. Regions high in PD, which are characterized by many distantly related taxa, were concentrated on the Southeast coast of Australia and throughout Tasmania (Fig 3b). These regions have been found to have a high diversity of conifers in previous studies (22). Fossils for 33 species of conifers have been found in North Western Tasmania, which indicates high conifer diversity relative to the size of the region (22). Tasmania and South Eastern Australia experienced a decline in conifer diversity after the early Oligocene (23) and other evidence suggests that most of the endemic genera, *Athrotaxis, Lagarostrobos* and *Microcachrys*, represent the only surviving members of lineages extending back to at least the earliest Cretaceous (24). These genera were also more geographically widespread in the past (23). *Athrotaxis, Lagarostrobos, Michrocachrys, Dislema* and *Phyllocladus* are largely restricted geographically to Tasmania. These findings suggest that these clades' ranges may be restricted by an ecological factor that has changed through time. Regions low in PD, which are characterized by many closely-related species, were more prevalent and were concentrated

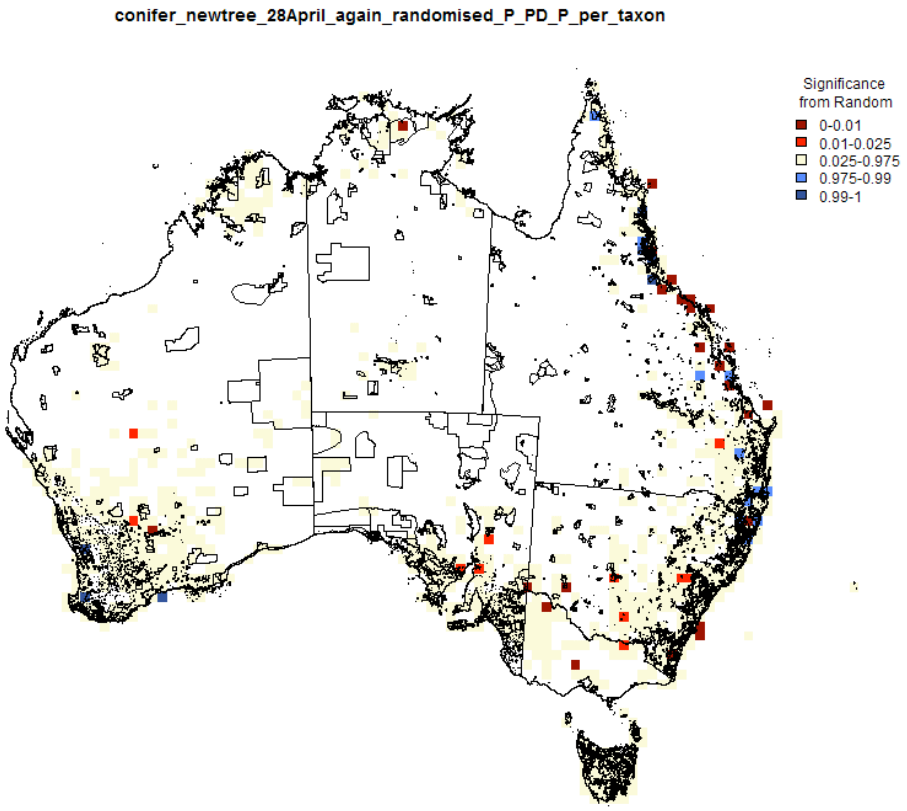



**Fig 5 Randomization of PD as a proportion of branch length overlaid with Australian Protected Regions (CAPAD 2010). Dark blue cells indicate regions significantly high in PD compared to a random distribution.**

inland of the coast and were primarily comprised of the genus *Callitris*, which is widespread throughout Australia (Fig 3b). Regions low in PD relative to their species richness estimate may be regions of isolated, large radiations (25).

WE and PE were strongly correlated ($r^2$= 0.87); however, WE underestimated the highest values of PE (Fig 4c). WE both overpredicted and underpredicted high PE scores (Fig 4c, due to the fact that closely related taxa may affect the result of PE if they contribute to the range of a clade with taxa in the study area (10). There were few regions high in PE, and they were concentrated Tasmania and on the Northeast coast, potentially due to the aforementioned endemic history of Tasmania.

### *Limitations and Future Directions*

A key limitation of my study is the spatial scale at which I performed analyses. Ecological and evolutionary patterns may differ at different spatial scales. Thus, it is important to re-analyze the data at different spatial scales, for instance 100,000$m^2$ grids or 25,000$m^2$ grids to check for consistency among the spatial scales. For this study, we chose 50,000$m^2$ grids because they have been shown to display subtleties of the data, and roughly estimate community sizes (16,17). Another spatial limitation stems from my use of the CAPAD 2010 shapefile in its entirety. This shapefile included all parklands, not only major reserves or conifer-specific reserves, and the number of vectors in this data layer made it difficult to interpret how effectively regions of high PD and PE are being conserved. Additionally, I was unable to answer one of my original research questions, which was to identify and map biogeographic regions that could be potential environmental explanations of PD/PE trends. I plan to continue this analysis and overlay these factors in the future.

Phylogenetically, my study is limited in its robustness, because I focused on a subset of species inhabiting the continent and this is a monophyletic group in relation to *Gingko biloba*, but polyphylys may be nested in these lineages. The phylogeny used for this study probably contains an error, a *matK* sequence for *Microstrobos niphophilus* which needs to be re-sequenced. Due to time constraints, I was unable to re-sequence it in time for this paper. It, however, does not affect the calculation of PD and PE as all of the branch lengths are incorporated that join sister taxa which share a spatial grid cell (10)

### *Broader Implications and Conclusions*

Examining the intersection of evolutionary history and spatial distribution of conifer species is a key method for properly informing conservation policy. Historically, approaches to biodiversity conservation have attempted to apply different concepts. Some have been more concerned with conserving rare species, while others have focused on key habitats. PD and PE are metrics that provide a way to account for both geography and evolutionary rarity. They are not in disagreement with SR and WE, instead they incorporate these metrics and provide more insight into the evolutionary and ecological processes that have occurred throughout time.

## Acknowledgements

This project would not have been possible without the support of Patina Mendez for her dedication, thorough editing, and constant positivity. Dr. Nathalie Nagalingum, Dr. Brent Mishler, Sonia Nosratinia and Nunzio Knerr were all integral to this project. They deserve so much more than a "thank you."

## References


1. Pimm, S. L.; G.J. Russell.; J.L. Gittleman,., and T.M. Brooks. 1995. The future of biodiversity. Science 269: 347–350.

2. Reaka-Kudla, M.L. 1997. Biodiversity II: understanding and protecting our biological resources. National Academy of Sciences, USA.

3. Lombard A. T., R. M. Cowling, R. L. Pressey, and A. G. Rebelo. 2003. Effectiveness of land classes as surrogates for species in conservation planning for the Cape Floristic Region. Biological Conservation 112:45-62.

4 Gaston and Fuller. 2009. The size of species' geographic ranges. Journal of Applied Ecology. 46: 1-9.

5. Crozier R. H. 1997. Preserving the Information Content of Species: Genetic Diversity, Phylogeny, and Conservation Worth. Annual Review of Ecology and Systematics 28:243-268.

6. Nixon K. C., Q. D. Wheeler. 1990. An amplification of the phylogenetic species concept. Cladistics 6:211-223.

7. Mishler B. D. 2009. Three Centuries of Paradigm Changes in Biological Classification: Is the End in Sight? Taxon 58:61-67.

8. Chao, A. 2005. Species richness estimation. Encyclopedia of Statistical Sciences. 7909-7916.

9. Crisp, M.D., S. Laffan, H. P. Linder, and A. Monro. 2001. Endemism in the Australian flora. Journal of Biogeography 29: 183-198.

10. Rosauer, D.F. and S.W. Laffan. 2008. Linking phylogenetic trees, taxonomy and geography to map phylogeography using Biodiverse. Taxonomic Data Working Group 2008. Perth, Australia.

11. Mooers, A. O. 2007. The diversity of biodiversity. Nature 445: 717-718.

12. Faith, D. P., and A. M. Baker. 2006. Phylogenetic diversity (PD) and biology conservation: some bioinformatics challenges. Evolutionary Bioinformatics 2: 70-77.

13. Meiri, S. and G. Mace. 2009. New taxonomy and the origin of species. Public Library of Science Biology 5: 1385- 1387.

14. Davies, J. T. and L. B. Buckley. 2011. Phylogenetic diversity as a window into the evolutionary and biogeographic histories of present- day richness gradients for mammals. The Royal Society 366: 2414-2425.

15. Mace, G. M., J.L. Gittleman, and A. Purvis. 2003. Preserving the tree of life. Science. 300: 1707- 1709.

16. Nagalingum, Nathalie. in prep. 2013

17. Mishler, B. D., N. Knerr, C. E. González-Orozco, A.H. Thornhill, S.W. Laffan, and J.T. Miller. (in review) Phylogenetic approaches to biodiversity, endemism, and conservation. Nature Communications

18. Ingelby, S. 2009. Endemism in Australian mammals. Australian Museum. http://australianmuseum.net.au/Endemism-in-Australian-mammals (Version 5/7/2012).

19. Leslie, A.B., J. M. Beaulieub, S.R. Hardeep, P.R. Cranea, M.J. Donoghueb, and S. Mathews. 2012. Hemisphere- scale differences in conifer evolutionary dynamics. Proceedings of the National Academy of Sciences of the United States of America. 109: 16217-16221.

20. Forest, F. R. Greyner, M. Rouget, J. Davies, R.M. Cowling, D.P. Faith, A. Blamford, J.C. Manning, S. Proches, M. van der Bank, G. Reeves, T.A. J. Hedderson, and V. Savolanien. 2007. Preserving the evolutionary potential of floras in biodiversity hotspots. Nature 445: 757-759

21. Rosauer, D., S. W. Laffan, M.D. Crisp., S.C. Donnellans, and L.G. Cook. 2009. Phylogenetic endemism: a new approach for identifying geographical concentrations of evolutionary history. Molecular Ecology 18: 4061-4072.

22. Jordan, . G.J. and R.S. Hill. 2002 Cenozoic plant macrofossil sites of Tasmania. Papers and Proceedings of the Royal Society of Tasmania, 136 .127-139

23. Hill, RS and Brodribb, TJ.1999. Turner Review No. 2 Southern Conifers in Time and Space. Australian Journal of Botany, 47: 639-696.

24. Quinn, C.J., R.A. Price, and P.A. Gadek. 2002. Familial concepts and relationships in the conifer based on rbcL and matK sequence comparisons. Kew Bulletin 57: 513- 531.

25. Fritz S. A., C. Rahbek. 2012. Global patterns of amphibian phylogenetic diversity. Journal of Biogeography 39:1373-1382.

## APPENDIX A: LIST OF INDIGENOUS CONIFER SPECIES

**List 1: Indigenous conifer species list**, including the outgroup used for this study, *Ginkgo biloba*

Actinostrobus arenarius
Callitris baileyi
Callitris columellaris
Callitris monticola
Callitris oblonga
Callitris roei
Microstrobos niphophilus
Agathis atropurpurea
Agathis microstachya
Agathis robusta
Araucaria bidwillii
Araucaria cunninghamii
Microcachrys tetragona
Actinostrobus acuminatus
Actinostrobus pyramidalis
Athrotaxis cupressoides
Athrotaxis selaginoides
Callitris canescens
Callitris drummondii
Callitris endlicheri
Callitris macleayana
Callitris muelleri
Callitris preissii
Callitris rhomboidea
Callitris verrucosa
Diselma archeri
Lagarostrobos franklinii
Microstrobos fitzgeraldii
Phyllocladus aspleniifolius
Podocarpus dispermus
Podocarpus drouynianus
Podocarpus elatus
Podocarpus grayae
Podocarpus lawrencei
Podocarpus smithii
Podocarpus spinulosus
Prumnopitys ladei
Sundacarpus amarus
Wollemia nobilis
Ginkgo biloba

## APPENDIX B: GENBANK ACCESSION NUMBERS

**Table 1: Genbank Accession numbers for matK and rbcL gene regions**

| rbcL | matK |
|---|---|
| JF725937 Actinostrobus arenarius | JF725837 Actinostrobus arenarius |
| EU161450 Actinostrobus pyramidalis | JF725831 Actinostrobus pyramidalis |
| AF502087 Agathis atropurpurea | EU025977 Agathis atropurpurea |
| AF508920 Agathis microstachya | EU025978 Agathis microstachya |
| EF490509 Agathis robusta | AF456371 Agathis robusta |
| AM920227 Araucaria bidwillii | EU025974 Araucaria bidwillii |
| EF490510 Araucaria cunninghamii | EU025975 Araucaria cunninghamii |
| JF725921 Athrotaxis cupressoides | JF725821 Athrotaxis cupressoides |
| JF725938 Athrotaxis selaginoides | JF725838 Athrotaxis selaginoides |
| JF725945 Callitris canescens | JF725845 Callitris canescens |
| JF725939 Callitris drummondii | JF725839 Callitris drummondii |
| JF725932 Callitris endlicheri | AY988331 Callitris endlicheri |
| JF725933 Callitris macleayana | JF725833 Callitris macleayana |
| JF725924 Callitris muelleri | JF725824 Callitris muelleri |
| JF725940 Callitris preissii | JF725840 Callitris preissii |
| L12537 Callitris rhomboidea | JF725825 Callitris rhomboidea |
| JF725942 Callitris verrucosa | JF725842 Callitris verrucosa |
| JF725926 Diselma archeri | JF725826 Diselma archeri |
| HM593609 Lagarostrobos franklinii | EU161486 Lagarostrobos franklinii |
| HM593611 Microcachrys tetragona | EU161483 Microcachrys tetragona |
| AF249646 Microstrobos fitzgeraldii | EU161484 Microstrobos fitzgeraldii |
| AF249647 Microstrobos niphophilus | |
| AF249651 Phyllocladus aspleniifolius | AY442147 Phyllocladus aspleniifolius |
| JF969685 Podocarpus dispermus | HM593741 Podocarpus dispermus |
| HM593639 Podocarpus drouynianus | HM593742 Podocarpus drouynianus |
| HM593641 Podocarpus elatus | HM593745 Podocarpus elatus |
| AF249608 Podocarpus grayae | HM593750 Podocarpus grayae |
| HM593651 Podocarpus lawrencii | HM593755 Podocarpus lawrencii |
| HM593675 Podocarpus smithii | HM593779 Podocarpus smithii |
| AF249630 Podocarpus spinulosus | HM593780 Podocarpus spinulosus |
| HM593620 Prumnopitys ladei | HM593723 Prumnopitys ladei |
| AF249663 Sundacarpus amarus | HM593788 Sundacarpus amarus |
| EF490508 Wollemia nobilis | AF456377 Wollemia nobilis |