**Title**

Adapting to extremes: reconstructing niche evolution in Australian *Acacia* over time and space.

Authors: Xia Hua[1,2*], Marcel Cardillo[1], Lindell Bromham[1]

[1]Macroevolution and Macroecology, Research School of Biology, Australian National University, Canberra ACT 0200 Australia

[2]Mathematical Sciences Institute, Australian National University, Canberra ACT 0200 Australia

**Abstract**

Macroevolutionary analysis is increasingly being used to study biodiversity responses to climate change, for example by using phylogenetic node ages to infer periods of diversification, or phylogenetic reconstruction of traits to infer adaptation to particular stresses. Here we apply a new macroevolutionary method to investigate the responses of a diverse plant genus, *Acacia*, to increasing aridity and salinity in Australia from the Miocene to the present. The Niche Evolution Model (NEMo) combines environmental niche modelling with phylogenetic comparative methods in a single statistical framework, to estimate current environmental tolerances, reconstruct the history of niche evolution and infer rates of change in key aspects of environmental tolerance. Using a large database of *Acacia* occurrence records and presence-absence survey sites, we find that both spatial and temporal patterns in niche evolution of *Acacia* are consistent with the aridification history of Australia and suggest high niche lability along both axes, which has allowed *Acacia* to quickly exploit new niches created during the aridification of the continent, and resulting in their current

dominance of many habitats across Australia. This study demonstrates that phylogenetic studies of niche evolution can move beyond application of simple trait-based models, allowing the underlying processes of speciation, adaptation and dispersal to be explicitly modelled in a macroecological and macroevolutionary context.

Keywords: aridity tolerance, salt tolerance, aridification, niche lability, Environmental Niche Modelling, Phylogenetic Comparative Methods.

## Introduction

The climatic and environmental history of Australia during the Neogene and Quaternary Periods has been a story of increasing aridification. Dry climates may have existed from the Eocene (Carpenter et al 2014), but evidence from charcoal and pollen records reveals an increase in fire frequency and prevalence of fire-adapted, scleromorphic plant taxa, pointing to an expansion of the arid zone from the early Miocene onwards (Byrne et al 2008). As the arid zone expanded, previously widespread mesic forest habitats contracted to the coastal margins, and plant groups responded in different ways (Byrne & Murphy 2020, Crisp et al 2004, Weston & Jordan 2017). Some groups (such as *Nothofagus*) remained confined to humid forest conditions and contracted with these habitats to small coastal refuges. Other groups responded to new ecological opportunities provided by the expanding arid zone, and diversified into numerous xeromorphic lineages. Some plant families, especially Proteaceae, Myrtaceae and Fabaceae, contain both mesic and xeromorphic lineages. Even some large genera, such as *Acacia, Grevillea* and *Hakea,* have both mesic and xeromorphic representatives. Plant groups such as these, in which close relatives occupy a range of ecoclimatic conditions, provide good case studies for studying the evolution of tolerance to extreme environmental conditions.

The expansion of the arid zone in Australia led to increasingly challenging environments characterized by stressful conditions for plant growth, including extreme drought, low soil fertility, frequent fires and high soil salinity. At the same time, areas that are extreme for most plants represent an opportunity for lineages to specialize, escape competition, and diversify. Arid-adapted lineages evolved a range of strategies to cope with water stress, broadly classified into escape, avoidance, and tolerance (Delzon 2015). We see all three strategies in *Acacia*, such as phyllode shedding (escape), scleromorphic structures like sunken stomata or needle-shaped phyllodes (avoidance), and physiological adaptations such as low leaf water potential (tolerance: Moore 2013). The wide range of mechanisms adopted by *Acacia* species to cope with water stress are reflected in their geographic distributions and varied climatic niches.

To understand how opportunities provided by the drying environments shaped present-day distribution and diversity, we need to examine the dynamics of lineage adaptation and diversification in response to extreme conditions. Macroevolutionary and macroecological analyses based on phylogenies have provided a variety of ways of investigating adaptation to harsh environments (Bromham et al. 2020). For example, the ages of nodes in molecular phylogenies have been used to infer bursts of diversification associated with periods of drying in *Acacia* (Renner et al. 2020), and the movement of rainforest-adapted lineages into dry environments (Crayn et al. 2006). In addition to relative timing of diversification, phylogenies can provide information on evolution of specific adaptations to extreme conditions. For example, Crisp et al. (2011) inferred the phylogenetic history of traits such as post-fire epicormic resprouting to investigate the timing of adaptation of *Eucalyptus* to increasing fire frequency. Jordan et al. (2008) reconstructed the evolution of stomatal protection structures (which are adaptations to dry climates) in Proteaceae to show that such structures evolved relatively few times in the ancestors of dry-climate clades. As

well as physiological and morphological traits, phylogenetic approaches have been applied to reconstructing the evolution of environmental niches. For example, by comparing alternative evolutionary models of climatic niche evolution for clades from different regions, Skeels and Cardillo (2017) showed that within four large genera (*Protea, Moraea, Banksia*, and *Hakea*), different lineages are evolving towards different climatic optima in different bioclimatic regions. Onstein et al. (2016) used the same approach to show that Proteaceae lineages in open vegetation and closed forest also have different inferred climatic optima.

Phylogenetic analyses such as these are useful, but have a number of limitations. To permit the application of standard stochastic models of trait evolution, such as Brownian Motion (BM) or Ornstein-Uhlenbeck (OU), in most macroevolutionary studies, the environmental niche is simplified to a point-estimate for a species, for example by taking the mean value across a species distribution (Evans et al 2009; Kozak & Wiens 2010; Münkemüller et al 2015; Renner et al 2020). But this is problematic for two main reasons. First, it averages away most of the potentially informative variation in environmental conditions across the range of a species. Second, models of trait evolution such as BM may not be appropriate for modelling the evolution of environmental tolerance. In particular, applying trait-evolution models to current environmental data to infer ancestral niches (e.g., Renner et al 2020, Skeels & Cardillo 2017) assumes that environmental traits such as aridity or temperature evolve randomly and independently along branches for each lineage. This assumption may often be violated because environmental niche evolution is likely to be driven by climatic changes, which most likely apply simultaneously and directionally across all lineages within a given region.

More fundamentally, phylogenetic analyses of niche evolution with changing climate compartmentalise the inference of contemporary niche (inferred from occurrence data) and the history of niche evolution (inferred from phylogenetic models). These are treated as two

separate problems, and subject to independent analyses with separate statistical frameworks. But contemporary niches are the product of macroevolutionary processes and inferring these processes requires accurate inference of the niche from contemporary distributions, so the two are fundamentally linked. The logical step, then, is to link their analyses too.

Here, we infer the history of adaptation to extreme environments by applying a novel method, the Niche Evolution Model (NEMo), that combines Environmental Niche Modelling (ENM) and reconstruction of niche evolution in a single statistical framework. This method uses occurrence records to characterise the distribution of each tip species in a phylogeny, and reconstructs niche evolution by explicitly modelling key driving processes: (1) shift in areas of tolerable conditions, (2) adaptation of lineages to changing conditions, (3) dispersal away from areas that are no longer suitable and into accessible areas of suitable habitat, (4) speciation, and (5) divergence along lineages' environmental tolerances. A strength of this method is that it allows us to infer the placement of these key processes on the phylogeny and to reconstruct ancestral niches at any time point, without making the unrealistic assumption that niches evolve under a simple stochastic trait-based model. Unlike previous approaches, NEMo can infer the occurrence density of these processes of niche evolution on the phylogeny, so it allows us to directly compare the timing of niche shift to periods of climatic change, or compare the timing of niche evolution along different environmental axes, using standard statistical tests, without pre-defining the covariance structure to account for phylogenetic nonindependence. Information on paleoclimate can either be incorporated into the model as prior information, or can be held aside as independent information against which to test the robustness of the inference.

We use NEMo to investigate macroevolutionary responses of the Australian plant genus *Acacia* to aridification. *Acacia* is one of the largest and most widely-distributed plant genera in Australia, with over 900 described species occupying nearly all major habitat types,

often as dominant taxa. Many *Acacia* species are known to be tolerant of a range of extreme environmental conditions, and several previous studies have used *Acacia* species as case studies for understanding tolerance to extremes, including salinity and other aspects of soil geochemistry (Bui et al. 2014a, 2014b), and water stress (Atkin et al. 2002). There has been rapid progress in understanding the evolutionary history of *Acacia* in recent years (Maslin et al. 2003; Murphy et al. 2010; Mishler et al. 2014). Nearly all *Acacia* species are conspicuous, woody plants, so spatial occurrence records for *Acacia* are abundant and widespread. We combine a curated database of *Acacia* occurrence records with extensive survey data, along with predictors of sampling bias, to characterise the current distribution of species in the phylogeny. We use NEMo to reconstruct the evolution of adaptation to increasing aridification in Australian *Acacia,* focusing on niche evolution in response to increased aridity and changes in the distribution and severity of soil salinity (frequently associated with aridification (Gilkes et al. 2003). Specifically, we ask two questions. First, have periods of rapid niche evolution along the aridity and salinity axes corresponded with major periods of aridification in Australia? Second, on which lineages did adaptation to increased aridity and salinity occur? Answering these two questions may help us understand whether *Acaci*a responded to continental aridification by a burst of adaptation, allowing the genus to radiate through the arid zone.

## Methods

### *1. Phylogenetic and Spatial Data*

We based our analysis on a published phylogeny of 508 *Acaci*a species, representing half of species currently recognized (Mishler et al. 2014). We timescaled branch lengths under a relaxed clock model using the R package ape (Paradis and Schliep 2018), with the root date

150 constrained between 21 and 23.9 million years ago (the estimated range of crown divergence
151 dates (Miller et al. 2013).

152 To describe the current environmental conditions under which each species is found,
153 we used a published dataset of 132,295 carefully checked presence locations of these 508
154 *Acacia* species, filtered from a total of 750,000 *Acacia* records from the Australian Virtual
155 Herbarium (Misher et al. 2014). We also assembled species presence-absence data from
156 floristic survey databases for New South Wales (BioNet Systematic Flora Survey;
157 environment.nsw.gov.au/research/VISplot.htm), Victoria (Victorian Biodiversity Atlas flora
158 records; environment.vic.gov.au/biodiversity/victorian-biodiversity-atlas), and Western
159 Australia (Flora Surveys of the Yilgarn; naturemap.dbca.wa.gov.au). We only included data
160 from complete floristic surveys conducted later than 2005 (to ensure consistency in
161 taxonomy) under government custodianship (to ensure consistent standards of survey
162 quality). As a result, our presence-absence survey data included 27,850 survey sites across
163 three states of Australia. Survey data are particularly valuable because they represent
164 presence/absence data, not just opportunistic presence records, and are not subject to the
165 same sampling biases as atlas or herbarium occurrence records. These presence-absence data
166 were analysed together with the presence locations. To represent degree of aridity and
167 salinity at each presence location for each species and each survey site, we used information
168 from the geographic layer of maximum monthly aridity index (Williams et al. 2012) and the
169 geographic layer of root zone soil electrical conductivity (Bui et al. 2017).

170 The majority of occurrence records for most species are presence-only records,
171 typically collected opportunistically, so their spatial distributions tend to be biased towards
172 regions that are more easily accessible to researchers, especially along road networks, and
173 areas of particular interest, such as protected areas or hotspots of diversity (Fithian et al.
174 2015; Haque et al. 2017). To correct for this sampling bias, we included two additional

factors: distance to the nearest road (using data from www.ga.gov.au/geoscience/national) and whether it is in a protected area, such as a conservation reserve or national park (from www.environment.gov.au/land/nrs/science/capad).

***2. Niche evolution model***

The Niche Evolution Model (NEMo) uses a Bayesian framework that includes three major components: model of niche evolution, phylogenetic comparative method (PCM), and ecological niche model (ENM). NEMo models species distributions in niche space, not geographic space. The aim is to infer not only species' niches at the contemporary tips of the phylogeny, but also to estimate the amount of change in niches along every branch of the phylogeny. This method has been shown to be effective at reconstructing past periods of niche evolution and better than standard ENMs at predicting aspects of contemporary niches (see Hua et al. submitted for details, available on bioRxiv).

The niche evolution component calculates how three aspects of niche (fundamental, available, realized) change under three basic processes (speciation, adaptation, dispersal). "Fundamental niche" represents the range of conditions that each species can tolerate at a given point in time, expressed as a range of values along a relevant environmental axis. We expect species to evolve along this environmental axis over evolutionary time, through the processes of adaptation to new conditions (by selecting for individuals at the extreme values of the distribution) and through speciation along environmental gradients (so that a species splits along its tolerance range leaving one daughter species that is more skewed toward the extreme conditions). "Available niche" describes the range of conditions in areas that are accessible to the species, regardless of whether the species can persist in those conditions. The potential of the species to occupy the available niche is governed by its fundamental niche and dispersal ability. We also include barriers to dispersal as nuisance parameters,

which could represent physical or biotic barriers. “Realized niche” consists of the locations where a species is found, which is shaped by the suitability of conditions and the opportunity to disperse to and occupy those areas.

The PCM part of the model infers the locations on the phylogeny of adaptation and speciation events, assuming these events are distributed along the branches of the phylogeny according to a Poisson process. Dispersal in niche space is modelled as a continuous process between any two events (adaptation or speciation), or nodes of the phylogeny. The distribution of these events on the phylogeny gives a possible history of niche evolution. Given this history, we use the model of niche evolution to calculate the three aspects of niche of each tip species. The ENM component of the method then assesses the fit of the inferred niches to the presence and/or absence locations of the tip species. The fit of the inferred niches to the data then feeds back to our inferences of the history of niche evolution and the niches of tip species. Details of the method, and description of tests of its performance and reliability, are given in Hua et al. (submitted), and the code with step-by-step instructions for implementing the model can be found at https://github.com/huaxia1985/NEMo. Below we briefly describe the parameters of the model applied to this case study.

***3. Model Parameters***

NEMo has two universal parameters which are assumed to be constant across the phylogeny: adaptation rate and dispersal rate. Additional parameters that are specific to each adaptation, speciation, and dispersal events are: 1) amount of environmental change in an adaptation event, 2) proportion of a tolerance range that is inherited by one descendent species (its sister species inherits the other part) from the ancestral species in a speciation event, 3) where a dispersal barrier locates along the environmental axis during a dispersal process.

We apply separate NEMo analyses to the evolution of environmental niches in *Acacia* along aridity and along salinity. For each of these niche axes, the values for the available niche and the fundamental niche at the root of the phylogeny were assumed to follow normal distributions. The starting values for their means and standard deviations were set to their maximum likelihood estimates given the likelihood of each tip species' presence-only data and presence-absence data, assuming that no events of niche evolution occurred on the phylogeny. The starting values for adaptation rate and dispersal rate were arbitrarily set to 0.03 and 1, respectively. Dispersal rate had rather flat prior, an exponential distribution with rate of 0.1. The parameters for the niches at the root and the adaptation rate had relatively narrow priors, a lognormal distribution with logarithmic mean of their initial values and logarithmic standard deviation of 0.5. The narrow priors for the root niches were used to ensure that root niches do not exceed observed maximum and minimum values (Williams et al. 2012; Bui et al. 2017) too often during MCMCs. The narrow prior for adaptation rate was used to ensure a finite likelihood value, as parameters of the fundamental niche change proportionally to adaptation rate on a log scale, so there is a narrow range of adaptation rate that gives finite likelihood values. The starting value for the occurrence rate of adaptation events and the occurrence rate of speciation events was arbitrarily set to 0.1. A narrow prior was applied on both occurrence rates, an exponential distribution with rate of 10 to reduce the degree of overestimation on the number of events (see Hua et al. submitted). For both aridity and salinity, we ran 20 independent MCMCs for $1\times10^6$ generations, with a thinning interval of 1000 and the first 600,000 generations discarded as burnin.

***4. Accounting for sampling bias***

We used an inhomogeneous Poisson point process (Renner et al. 2015) as the ENM component of NEMo. Because opportunistic sampling bias can influence the niche estimation

(Fithian et al. 2015), we modified the likelihood function to account for sampling bias, using the method of Fithian et al. (2015). The log-likelihood of the presence-only data of each tip species is modified as,

$$\ell_P = \sum_{i \in I_{PO}} \left[\log(\lambda_{r_i}) + \delta z_i\right] - \frac{\mathcal{D}}{N} \sum_{i \in I_{BG}} \lambda_{f_i} \lambda_{a_i} e^{\delta z_i}$$

, where $z_i$ are values of the factors that influence sampling effort at the $i^{\text{th}}$ location and $\delta$ are the coefficients of the two sampling factors we included - distance of the location to the nearest road and whether the location is in a protected area - which are assumed to have the same value for all tip species. The starting values of coefficients were set to their maximum likelihood estimates given the likelihood function, and a rather flat prior was used: an exponential distribution with rate of 0.1. In the second term of the equation, $\lambda_f$ is the inferred fundamental niche of a tip species, giving the proportion of individuals of the species that can tolerate a certain level of aridity and salinity, and $\lambda_a$ is the inferred available niche of the tip species, giving the proportion of areas accessible to the species that have certain level of aridity and salinity. If we were not accounting for sampling bias, the second term of the equation would be a direct integral of $\lambda_f \lambda_a$ (Hua et al. submitted), but because each location now has a specific value for sampling bias, we can only approximate the integral by randomly sampling $N$ locations from the surroundings of the species presence locations to construct a set of background locations $I_{BG}$ and averaging the term to integrate over all the background locations. Note that these background locations are only used to numerically approximate the second term of the equation, so they are not pseudo-absence locations (Fithian et al. 2015).

For each *Acacia* species, background locations were sampled uniformly from areas within an arbitrary large radius (20 *km*) of each presence location of the species, so the weight of each background location is $\frac{\mathcal{D}}{N}$, where $\mathcal{D}$ is the total area that falls within the 20 *km*

radius of each species presence location, and $N$ is the total number of background locations, which was ten times the number of the species presence locations, with a minimum of 1000 for species with few presence locations. Based on the environmental conditions over all the background locations, we used Sturges' formula (Sturges 1926) to decide the bin size to discretise the aridity index and the electrical conductivity.

***5. Estimating rate and amount of niche evolution***

The amount of niche evolution per branch is calculated as the sum of the differences in the mean of the realized niche before and after each event (both adaptation and speciation events) inferred along each branch of the phylogeny over the posterior samples, calculating both the net amount (positive or negative) and the absolute amount. The net amount of niche evolution divided by branch length gives the direction of niche evolution, and the absolute amount of niche evolution divided by branch length gives the rate of niche evolution. Last, we calculated the intensity of the inferred environmental change during adaptation events at each time point over the whole phylogeny and compared the distribution of this intensity over time to the aridification history of Australia (see details in Hua et al. submitted).

## Results

***1. Spatial Patterns of Niche Evolution***

The *Acacia* phylogeny is characterized by a mixture of large, young clades (recent, rapid divergences), and smaller, older clades (longer-persisting lineages: Figure 1). Three recent radiations, each around 1 Ma old, include most present-day species in Northern Australia, Southern Australia, and Eastern Australia. Of these, the Northern Australian clade has the highest average rate of niche evolution on both aridity and salinity axes (Figure 1). The smaller, older clades are mainly in Western Australia, with many extant species originating

earlier than the three large radiations. These smaller older clades have, on average, much lower rate of niche evolution on both aridity and salinity axes than the recent radiations (Figure 1). Present-day species that occupy regions that are arid and prone to high salinity (deserts and Mediterranean ecosystems) are found throughout the phylogeny, forming small clusters of related species. In contrast, present-day species that occupy the mesic zone are found mainly in young clades as large clusters of closely-related species.

The net amount of niche evolution inferred by NEMo along the phylogeny is shown in Figure 1. Most of the evolutionary change on both aridity and salinity axes has been pushing species towards tolerance of the extreme conditions. We can see this in the patterns of niche evolution along the aridity axis in different habitat types (Figure 2A). The largest changes in niche are in lineages currently found in the most arid environments, while little or no niche shift is in lineages currently distributed in forests. The same pattern is repeated over the whole phylogeny, and also within each of the recent radiations (Figure 2A). Similarly, on the salinity axis, NEMo inferred the greatest amount of niche evolution towards higher salinity leading to species currently distributed in the Mediterranean ecosystems (Figure 2B). The same pattern is inferred in the three recent radiations, where there is little change along the salinity axis in lineages currently distributed in forests (Figure 2B).

***2. Temporal Patterns of Niche Evolution***

On the aridity axis, the inferred intensity of environmental change during each adaptation event peaks at 13 Ma and near the present (after 2.5 Ma), corresponding to major drying periods of Australia (Figure 3). On the salinity axis, the inferred intensity of environmental change during each adaptation event peaks near the present (after 2.5 Ma), with two additional lower peaks around 10 Ma and 17 Ma (Figure 3).

Across the phylogeny, the net amount of inferred niche evolution due to both adaptation and speciation events along the salinity and aridity axes showed a significant positive correlation (Figure 4; Kendall's rank correlation coefficient=0.25, $p<0.001$). Although significant, the correlation is mostly driven by niche evolution along the terminal branches, and there is a great deal of scatter, with relatively more aridity evolution on tip branches leading to present-day species found in the arid zone (Deserts & Xeric Shrublands), and relatively more salinity evolution on tip branches leading to present-day species found in Mediterranean ecosystems. The greater amounts of aridity and salinity evolution along the tip branches are not due to longer branch length, as NEMo also infers large amount of aridity and salinity evolution along very short tip branches (Figure 4).

## Discussion

Aridification has been the prevailing environmental trend across most of the Australian continent through the Neogene and Quaternary. In order to understand the history underlying Australia's current plant diversity, we need effective methods for reconstructing the evolutionary responses of plant clades to the environmental changes brought about by aridification. In this study, we use the large, widespread genus *Acacia* as a case study to reconstruct its history of niche evolution along aridity and salinity axes, using the newly developed method NEMo. We reconstruct both spatial and temporal patterns in niche evolution of *Acacia* and show (1) a preponderance of evolutionary change in niches towards more extreme values of aridity and salinity, (2) a concentration of niche evolution in the ancestors of species currently found in arid and Mediterranean regions in both the old clades and the recent radiations, (3) peaks in niche evolution throughout the aridification history of Australia, (4) correlated evolution of aridity and salinity tolerance, particularly along tip branches. As we explain below, these results are consistent with the adaptation of *Acacia*

lineages to drier climates, allowing them to exploit new niches opening up through the aridification of the continent. Our results point to an evolutionary flexibility in *Acacia* that is likely to have played a key role in their current near-ubiquitous distribution and dominance of many habitats across Australia.

The preponderance of niche evolution towards more extreme values of both aridity and salinity suggests that aridification has been a dominant driver of *Acacia* evolution and diversification across Australia during the Neogene & Quaternary. The concentration of niche evolution on branches leading to species currently found in more arid and salt-affected regions, in both old and recent clades, suggests that the recent radiations of *Acacia* were not triggered by adaptation to aridity and salinity. Instead, extant species occupying arid and salt-affected regions are likely to have evolved from ancestral lineages in mesic areas which do not show high rates of niche evolution along these axes. This interpretation indicates evolutionary flexibility in *Acacia*, which is supported by the high rate of niche evolution along both aridity and salinity axes in the three recent radiations. This enhanced rate of niche evolution was also found by a recent study of diversification in *Acacia* using traditional PCMs, which hypothesized that Pleistocene glacial-interglacial climate cycling (< 2.5Ma) drove rapid climatic adaptation, which promoted diversification and led to large radiations (Renner et al. 2020).

Unlike traditional PCMs, NEMo is able to identify key events of niche evolution along the *Acacia* phylogeny. We found that most changes in niche along the aridity axis were in the recent radiations, predominantly after 1Ma and along tip branches. But niche evolution in aridity was also inferred along two branches ancestral to all these radiations (marked with asterisk in Figure 1), between 12 to 15 Ma and around 5Ma (Figure 1). It is likely that the niche evolution events that occurred along the two ancestral branches left descendant lineages of *Acacia* able to tolerate higher aridity than the other lineages in Tropical and

374 Temperate forests and grasslands during the Pleistocene, so that *Acacia* was able to replace
375 many other groups during the drying period of the Pleistocene. These two pulses of niche
376 evolution along the aridity axis are also consistent with the aridification history of Australia.
377 *Acacia* emerged during a period of aridification (28-23Ma), and we found peaks of
378 adaptation to increasing aridity in *Acacia* during the two subsequent drying periods,14-5Ma
379 and after 5Ma (Byrne et al. 2008).

380 Comparing salt tolerance to aridity tolerance in *Acacia*, we found that the amount of
381 niche evolution along aridity and salinity axes was highly correlated across branches of the
382 *Acacia* phylogeny. This result is consistent with the observation that saline soils are
383 frequently associated with aridity, so many species are adapted to both aridity and salinity
384 (Gong et al. 2017). Correlated evolution was also suggested between salinity and alkalinity
385 tolerance in *Acacia* using phylogenetic independent contrasts (Bui et al. 2014b). The NEMo
386 method offers a way to account for phylogenetic nonindependence in testing correlated
387 evolution, without making *a priori* assumptions about the phylogenetic covariance matrix. In
388 addition, the NEMo method allows us to detect branches that have large amounts of niche
389 evolution along both aridity and salinity axes. These branches are mostly terminal branches
390 scattered across the *Acacia* phylogeny and are mostly ancestral to species currently
391 distributed in areas with both high aridity and high salinity. This result is consistent with high
392 niche lability along both aridity and salinity in Australian *Acacia*, as suggested by the high
393 rate of niche evolution in the three radiations.

394 Inferred peaks of increased salinity are on recent lineages, which is consistent with
395 previous findings of a phylogenetically "tippy" distribution of salt tolerant species in grasses
396 (Bennett et al. 2013; Moray et al. 2015). Consequently, even if salt tolerance has evolved
397 many times through old adaptation events, these events rarely gave rise to persistently salt
398 tolerant lineages of large clades of related salt tolerant taxa, leaving less phylogenetic signal

for us to reconstruct adaptation events in the deep part of the phylogeny than adaptation events near the tip of the phylogeny (Bromham 2016). The high niche lability in Australian *Acacia* further agrees with previous hypothesis that the tippy distribution of salt tolerant species is due to high niche lability (Bromham et al. 2016).

It is important to note that for the NEMo analysis, as for most macroevolutionary analyses, we have assumed that the phylogeny has accurate topology and provides an approximate scaling of evolutionary events in time. Clearly, phylogenetic estimates of either topology or the timing of evolutionary events are not error-free (Bromham et al. 2018; Guindon 2020; Nie et al. 2019). In particular, uncertainty or bias in branch length estimates arising from imperfect characterisation of molecular rates over time may have a non-trivial impact on macroevolutionary inference (Duchêne et al. 2017). Phylogenies such as this one where the only calibration is a root date may be particularly vulnerable to errors in position of node heights in some parts of the tree (Duchêne et al. 2014). One possible solution is to apply the NEMo analysis on posterior samples of phylogenies, but perhaps a better advance will be to embed the NEMo method into phylogenetic reconstruction, which allows us to directly link niche evolution events to diversification events (Hua et al. submitted). Another issue with our current inference of niche evolution in *Aacia* is incomplete sampling, as the *Aacia* phylogeny only includes half of the known extant species. One obvious solution is to sample more species, but another way to increase the amount of information in the analysis of niche evolution is to incorporate historical data, such as paleoclimate, as priors for the occurrence density of niche evolution events, to supplement current species distribution data (Hua et al. submitted). Here we have used only part of the available information in the inference process in order to demonstrate the utility of the method (so we can compare our inferred evolutionary history to known paleoclimatic history), but the method allows for all available information to contribute to stronger inference.

424

425 **References**

426 Atkin O., Schortemeyer M., McFarlane N., Evans J.R. 2002. Variation in the components of
427 relative growth rate in 10 *Acacia* species from contrasting environments. Plant Cell
428 Environ. 21:1007–1017.
429 Bennett T.H., Flowers T.J., Bromham L. 2013. Repeated evolution of salt-tolerance in
430 grasses. Biol. Lett. 9:20130029.
431 Bromham L. 2016. Testing hypotheses in macroevolution. Stud. Hist. Philos. Sci. A 55:47–
432 59.
433 Bromham L., Hua X., Cardillo M. 2016. Detecting macroevolutionary self-destruction from
434 phylogenies. Syst. Biol. 65:109–127.
435 Bromham L., Duchêne S., Hua X., Ritchie A., Duchêne D., Ho S. 2018. Bayesian molecular
436 dating: opening up the black box. Biol. Rev. 93:1165–1191.
437 Bromham L., Hua X., Cardillo M. 2020. Macroevolutionary and macroecological approaches
438 to understanding the evolution of stress tolerance in plants. Plant Cell Environ. 43:1–
439 15.
440 Byrne M., Yeates D.K., Joseph L., Kearney M., Bowler J., Williams M.A., Cooper S.,
441 Donnellan S.C., Keogh J.S., Leys R., Melville J., Murphy D.J., Porch N., Wyrwoll
442 K.H. 2008. Birth of a biome: insights into the assembly and maintenance of the
443 Australian arid zone biota. Mol. Ecol. 17:4398–4417.
444 Byrne M., Murphy D.J. 2020. The origins and evolutionary history of xerophytic vegetation
445 in Australia. Aust. J. Bot. 68:195–207.
446 Bui E.N., González-Orozco C.E., Miller J.T. 2014a. *Acacia*, climate, and geochemistry in
447 Australia. Plant and Soil 281:161–175.

448 Bui E.N., Thornhill A., Miller J.T. 2014b. Salt- and alkaline-tolerance are linked in *Acacia*.
449 Biol. Lett. 10:20140278.
450 Bui E.N., Wilford J., de Caritat P. 2017. Predictive spatial modelling for mapping soil salinity
451 at continental scale, EGU General Assembly, Vienna, Austria. doi:
452 10.13140/RG.2.2.25552.20483.
453 Carpenter R.J., McLoughlin S., Hill R.S., McNamara K.J., Jordan G.J. 2014. Early evidence
454 of xeromorphy in angiosperms: stomatal encryption in a new eocene species of
455 *Banksia* (Proteaceae) from Western Australia. Am. J. Bot. 101:1486–1497.
456 Crayn D.M., Rossetto M., Maynard D.J. 2006. Molecular phylogeny and dating reveals an
457 Oligo-Miocene radiation of dry-adapted shrubs (former Tremandraceae) from
458 rainforest tree progenitors (Elaeocarpaceae) in Australia. Am. J. Bot. 93:1328–1342.
459 Crisp M., Cook L., Steane D.A. 2004. Radiation of the Australian flora: what can
460 comparisons of molecular phylogenies across multiple taxa tell us about the evolution
461 of diversity in present day communities? Philos. Trans. R. Soc. B 359:1551–1571.
462 Crisp M.D., Burrows G.E., Cook L.G., Thornhill A.H., Bowman D.M.J.S. 2011. Flammable
463 biomes dominated by eucalypts originated at the Cretaceous–Palaeogene boundary.
464 Nat. Commun. 2:193.
465 Delzon S. 2015. New insight into leaf drought tolerance. Funct. Ecol. 29:1247–1249.
466 Duchêne S., Lanfear R., Ho S.Y.W. 2014. The impact of calibration and clock-model choice
467 on molecular estimates of divergence times. Mol. Phylogenetics Evol. 78:277–289.
468 Duchêne D.A., Hua X., Bromham L. 2017. Phylogenetic estimates of diversification rate are
469 affected by molecular rate variation. Journal of Evolutionary Biology 30:1884–1897.
470 Evans M.E.K., Smith S.A., Flynn R.S., Donoghue M.J. 2009. Climate, niche evolution, and
471 diversification of the "bird-cage" evening primroses (Oenothera, sections Anogra and
472 Kleinia). Am. Nat. 173:225–240.

473 Fithian W., Elith J., Hastie T., Keith D.A. 2015. Bias correction in species distribution
474 models: pooling survey and collection data for multiple species. Methods Ecol. Evol.
475 6:424–438.
476 Gilkes B., Lee S., Singh B. 2003. The imprinting of aridity upon a lateritic landscape: an
477 illustration from southwestern Australia. C. R. Geosci. 335:1207–1218.
478 Gong Y., Lv G., Guo Z., Chen Y., Cao J. 2017. Influence of aridity and salinity on plant
479 nutrients scales up from species to community level in a desert ecosystem. Sci. Rep.
480 7:6811.
481 Guindon S. 2020. Rates and Rocks: Strengths and Weaknesses of Molecular Dating Methods.
482 Front. Genet. 11:526–526.
483 Haque M.M., Nipperess D.A., Gallagher R.V., Beaumont L.J. 2017. How well documented is
484 Australia's flora? Understanding spatial bias in vouchered plant specimens. Austral
485 Ecol. 42:690–699.
486 Hua X., Cardillo M., Bromham L. 2020. Tracking niche change through time: simultaneous
487 inference of ecological niche evolution and estimation of contemporary niches.
488 bioRxiv XXXXXXXXXXX
489 Jordan G. J., Weston P.H., Carpenter R.J., Dillon R.A., Brodribb T.J. 2008. The evolutionary
490 relations of sunken, covered, and encrypted stomata to dry habitats in Proteaceae. Am.
491 J. Bot. 95:521–530.
492 Kozak K., Wiens J. 2010. Accelerated rates of climatic-niche evolution underlie rapid species
493 diversification. Ecol. Lett. 13:1378–1389.
494 Maslin B.R., Miller J.T., Seigler D.S. 2003. Overview of the generic status of *Acacia*
495 (Leguminosae: Mimosoideae). Aust. Syst. Bot. 16:1–18.

496 Miller J.T., Murphy, D.J., Brown G.K., Richardson D.M., González-Orozco C.E. 2011. The
497 evolution and phylogenetic placement of invasive Australian *Acacia* species.
498 Diversity Distrib. 17:848–860.
499 Miller J.T., Murphy D.J., Ho S.Y.W., Cantrill D.J., Seigler D. 2013. Comparative dating of
500 *Acacia*: combining fossils and multiple phylogenies to infer ages of clades with poor
501 fossil records. Aust. J. Bot. 61:436−445.
502 Mishler B.D., Knerr N., González-Orozco C.E., Thornhill A.H., Laffan S.W., Miller J.T.
503 2014. Phylogenetic measures of biodiversity and neo- and paleo-endemism in
504 Australian *Acacia*. Nat. Comm. 5:4473.
505 Moore G.M. 2013. Adaptations of Australian tree species relevant to water scarcity in the
506 urban forest. Arboric. Urban. For. 39:109−115.
507 Moray C., Hua X., Bromham L. 2015. Salt tolerance is evolutionarily labile in a diverse set
508 of angiosperm families. BMC Evol. Biol. 15:90.
509 Münkemüller, T., Boucher, F.C., Thuiller, W., & Lavergne, S. (2015). Phylogenetic niche
510 conservatism – common pitfalls and ways forward. Funct. Ecol. 29:627−639.
511 Murphy D. J., Brown G. K., Miller J. T., Ladiges P. Y. 2010. Molecular phylogeny of *Acacia*
512 Mill. (Mimosoideae: Leguminosae): evidence for major clades and informal
513 classification. Taxon 59:7–19.
514 Nie Y., Foster C.S.P., Zhu T., Yao R., Duchêne D.A., Ho S.Y.W., Zhong B. 2019.
515 Accounting for uncertainty in the evolutionary timescale of green plants through
516 clock-partitioning and fossil calibration strategies. Syst. Biol. 69:1−16.
517 Onstein R.E., Jordan G.J., Sauquet H., Weston P.H., Bouchenak-Khelladi Y., Carpenter R.J.,
518 Linder H.P. 2016. Evolutionary radiations of Proteaceae are triggered by the
519 interaction between traits and climates in open habitats. Glob. Ecol. Biogeogr.
520 25:1239–1251.

Owen C.L., Marshall D.C., Hill K.B.R., Simon C. 2017. How the aridification of Australia structured the biogeography and influenced the diversification of a large lineage of Australian cicadas. Syst. Biol. 66:569−589.

Paradis E., Schliep K. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics, bty633.

Renner I.W., Elith J., Baddeley A., Fithian W., Hastie T., Phillips S.J., Popovic G., Warton D.I. 2015. Point process models for presence-only analysis. Methods Ecol. Evol. 6:366–379.

Renner M.A.M., Foster C.S.P., Miller J.T., Murphy D.J. 2020. Increased diversification rates are coupled with higher rates of climate space exploration in Australian *Acacia* (Mimosoideae). New Phytol. 226:609–622.

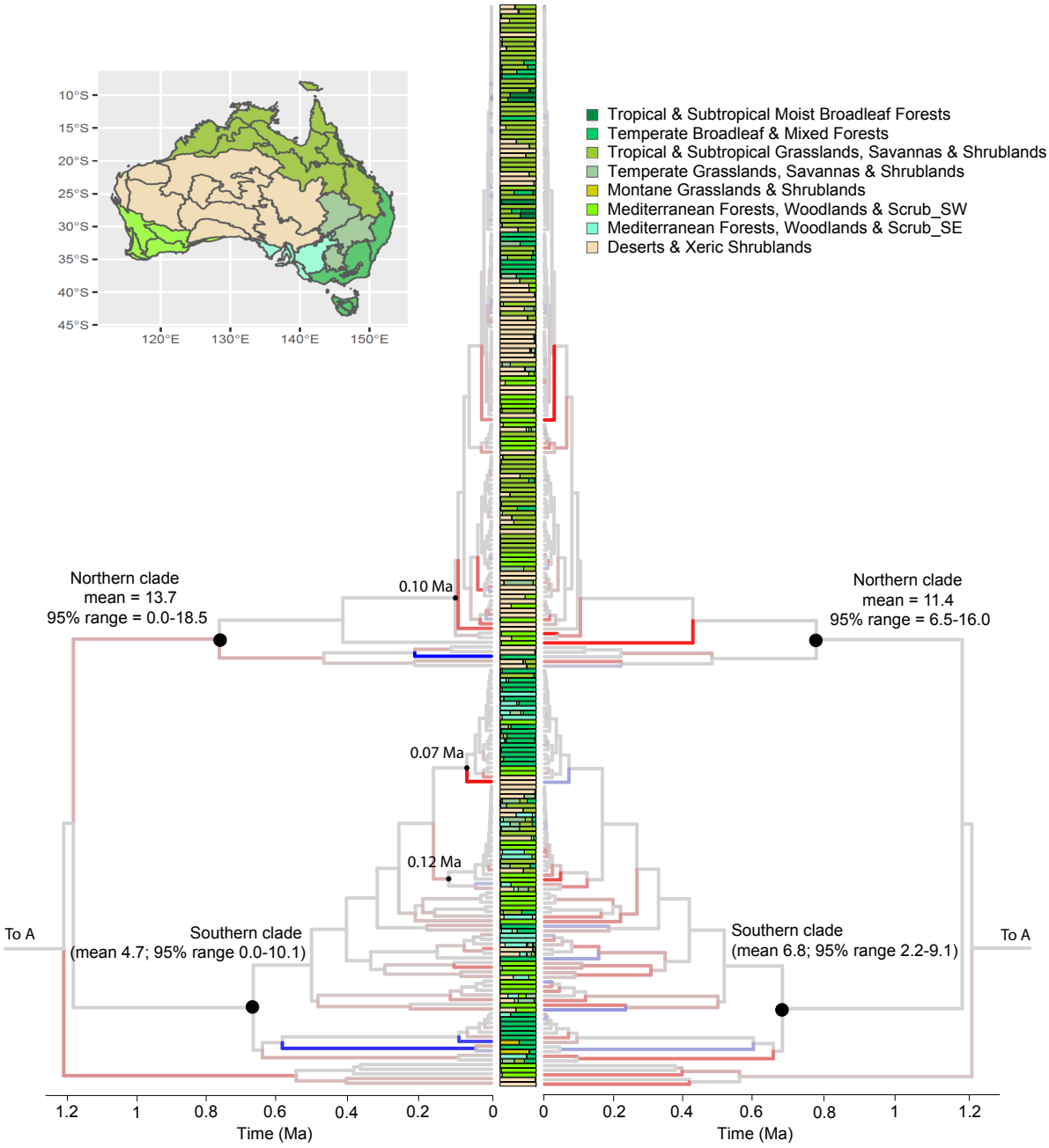
Skeels A., Cardillo M. 2017. Environmental niche conservatism explains the accumulation of species richness in Mediterranean-hotspot plant genera. Evolution 71: 582–594.

Sturges H.A. 1926. The choice of a class interval. J. Am. Stat. Assoc. 21:65–66.

Weston P.H., Jordan G.J. 2017. Evolutionary biogeography of the Australian flora in the Cenozoic era. In 'Australian vegetation,' 3rd edn. (Ed. DA Keith) pp. 40–62. (Cambridge University Press: Cambridge, UK)

Williams K.J., Belbin L., Austin M.P., Stein J.L., Ferrier S. 2012. Which environmental variables should I use in my biodiversity model? Int. J. Geogr. Inf. Sci. 26:2009–2047.

Figure 1. Phylogenetic distribution of the inferred amount of niche evolution along axes of aridity (left) and salinity (right). The blue to red color scale gives the sum of the amount of change in the mean of realized niche values before and after each niche evolution event inferred along each branch, averaged over posterior samples. Shifts towards less extreme conditions are in blue and shifts towards more extreme conditions are in red. The ecoregion map shows the current distribution of eight ecoregions in Australia. Each tip species has a stacked bar showing the relative proportion of its current presence locations in each ecoregion, in the same color scheme as the ecoregion map. We also include two additional maps. Left one shows the spatial variation in the aridity index in Australia, with lower values in more arid areas. Right one shows the variation in root zone soil electrical conductivity, with higher values in areas with higher salinity. The mean and the 95% range of estimated rate of niche evolution over all posterior samples are marked for each major clade in the phylogeny.


10°S
15°S
20°S
25°S
30°S
35°S
40°S
45°S
120°E
130°E
140°E
150°E
Tropical & Subtropical Moist Broadleaf Forests
Temperate Broadleaf & Mixed Forests
Tropical & Subtropical Grasslands, Savannas & Shrublands
Temperate Grasslands, Savannas & Shrublands
Montane Grasslands & Shrublands
Mediterranean Forests, Woodlands & Scrub_SW
Mediterranean Forests, Woodlands & Scrub_SE
Deserts & Xeric Shrublands
Northern clade
mean = 13.7
95% range = 0.0-18.5
0.10 Ma
Northern clade
mean = 11.4
95% range = 6.5-16.0
0.07 Ma
0.12 Ma
To A
Southern clade
(mean 4.7; 95% range 0.0-10.1)
Southern clade
(mean 6.8; 95% range 2.2-9.1)
To A
1.2
1
0.8
0.6
0.4
0.2
0
Time (Ma)
0
0.2
0.4
0.6
0.8
1
1.2
Time (Ma)

556

557

Figure 1 cont.

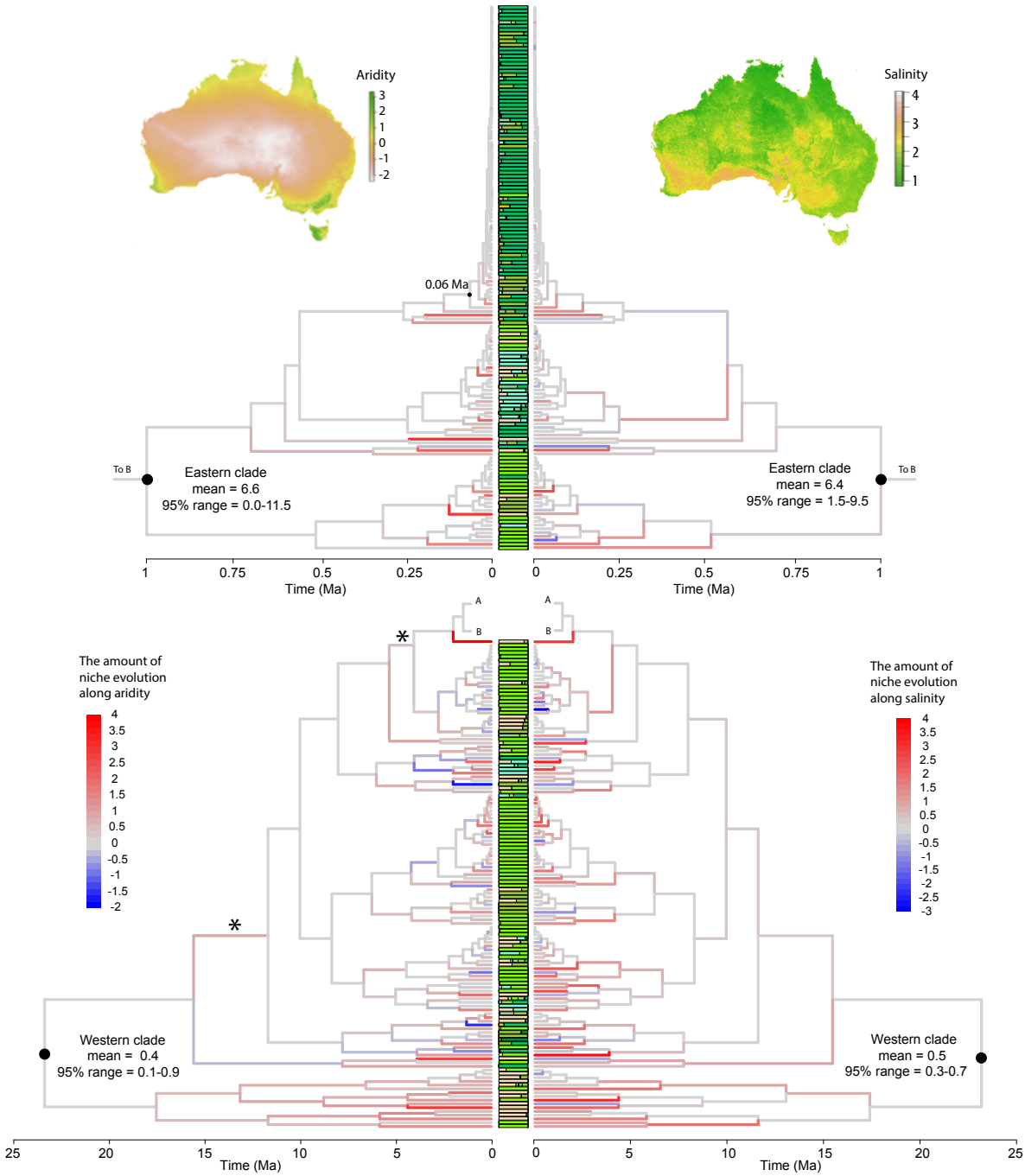
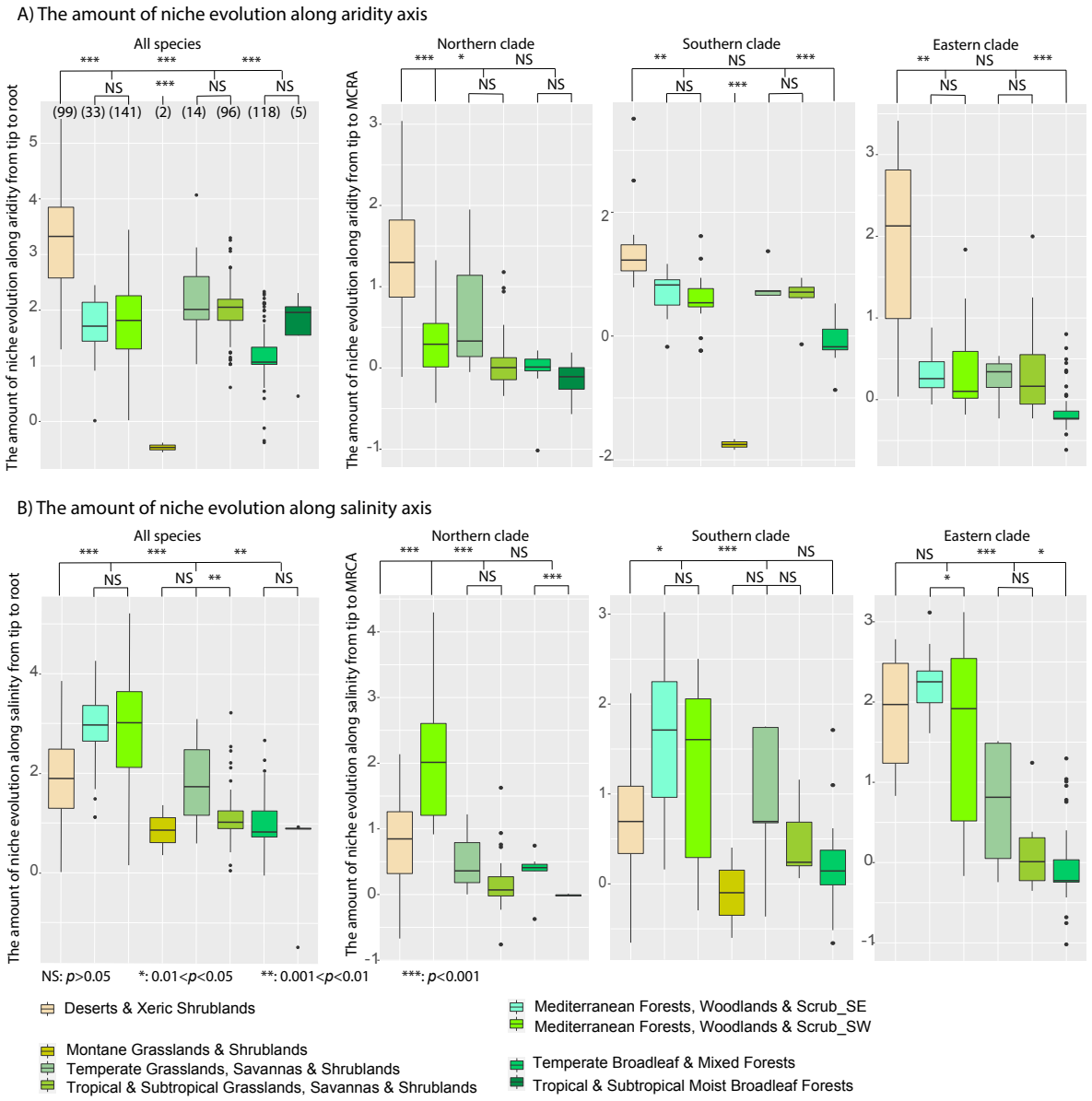
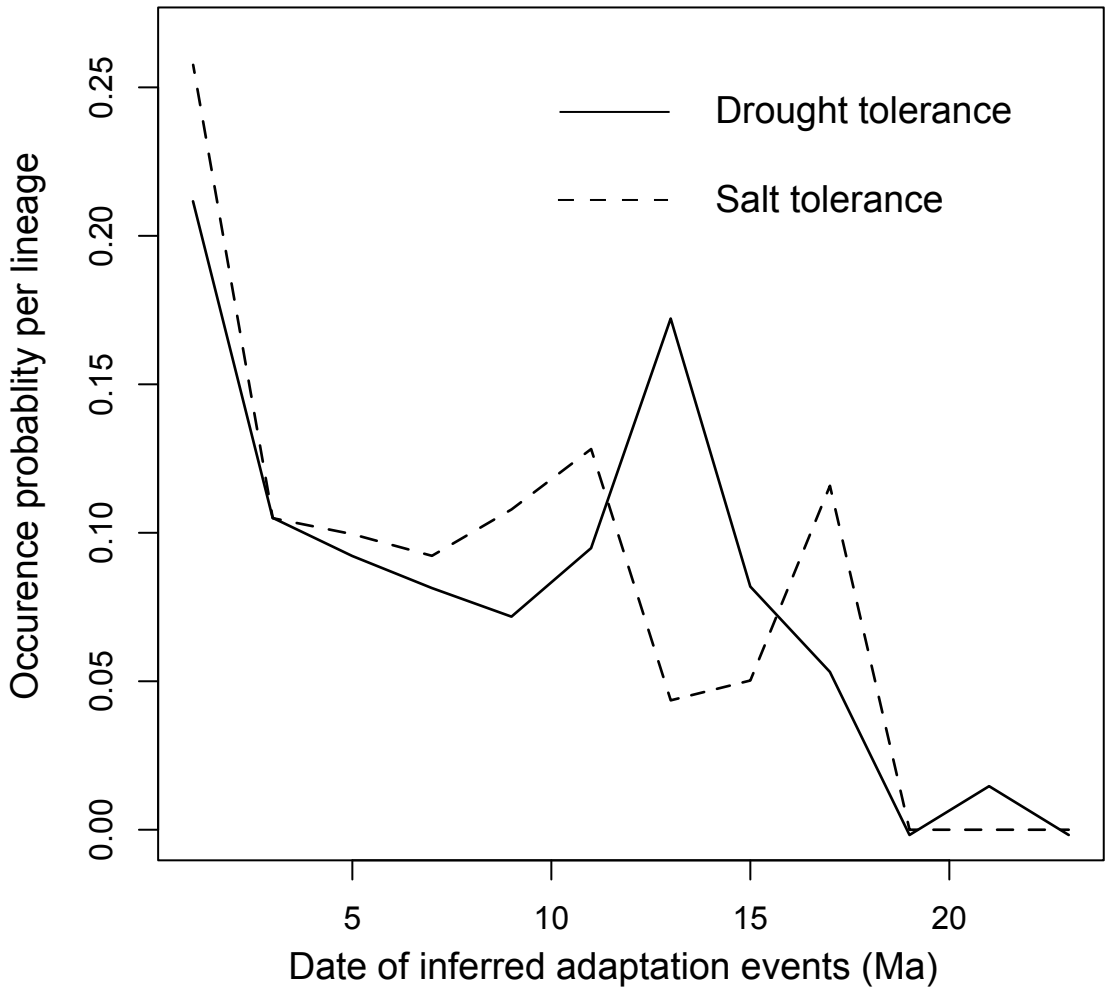


Figure 2. Boxplots of the net amount of niche evolution along aridity axis A) and salinity axis B). “All species” plots the sum of the amount of niche evolution along the branches that connect the root to each tip of the phylogeny. “Northern clade”, “Southern clade”, “Eastern clade” plot the sum of the amount of niche evolution along the branches that connect the MCRA of the three clades of recent radiation to each tip in these clades in Figure 1. Tips are grouped by the ecoregion that the majority of current presence location of the tip species fall in, using the same color scheme as the ecoregion map in Figure 1. T-test is used to first compare within vegetation types: desert, Mediterranean, grasslands, forests. If regions of the same vegetation type do not differ, then these regions are grouped and compared to other vegetation types. The statistical significance of a t-test between two bioregions linked by lines is shown above the corresponding horizontal line.


A) The amount of niche evolution along aridity axis
All species
Northern clade
Southern clade
Eastern clade
The amount of niche evolution along aridity from tip to root
The amount of niche evolution along aridity from tip to MCRA
(99) (33) (141) (2) (14) (96) (118) (5)
B) The amount of niche evolution along salinity axis
The amount of niche evolution along salinity from tip to root
The amount of niche evolution along salinity from tip to MRCA
NS: p>0.05 *: 0.01<p<0.05 **: 0.001<p<0.01 ***: p<0.001
Deserts & Xeric Shrublands
Montane Grasslands & Shrublands
Temperate Grasslands, Savannas & Shrublands
Tropical & Subtropical Grasslands, Savannas & Shrublands
Mediterranean Forests, Woodlands & Scrub_SE
Mediterranean Forests, Woodlands & Scrub_SW
Temperate Broadleaf & Mixed Forests
Tropical & Subtropical Moist Broadleaf Forests

574 Figure 3. Histogram of the inferred dates of adaptation events in *Acacia* drought tolerance

575 (solid) and salt tolerance (dashed).




576

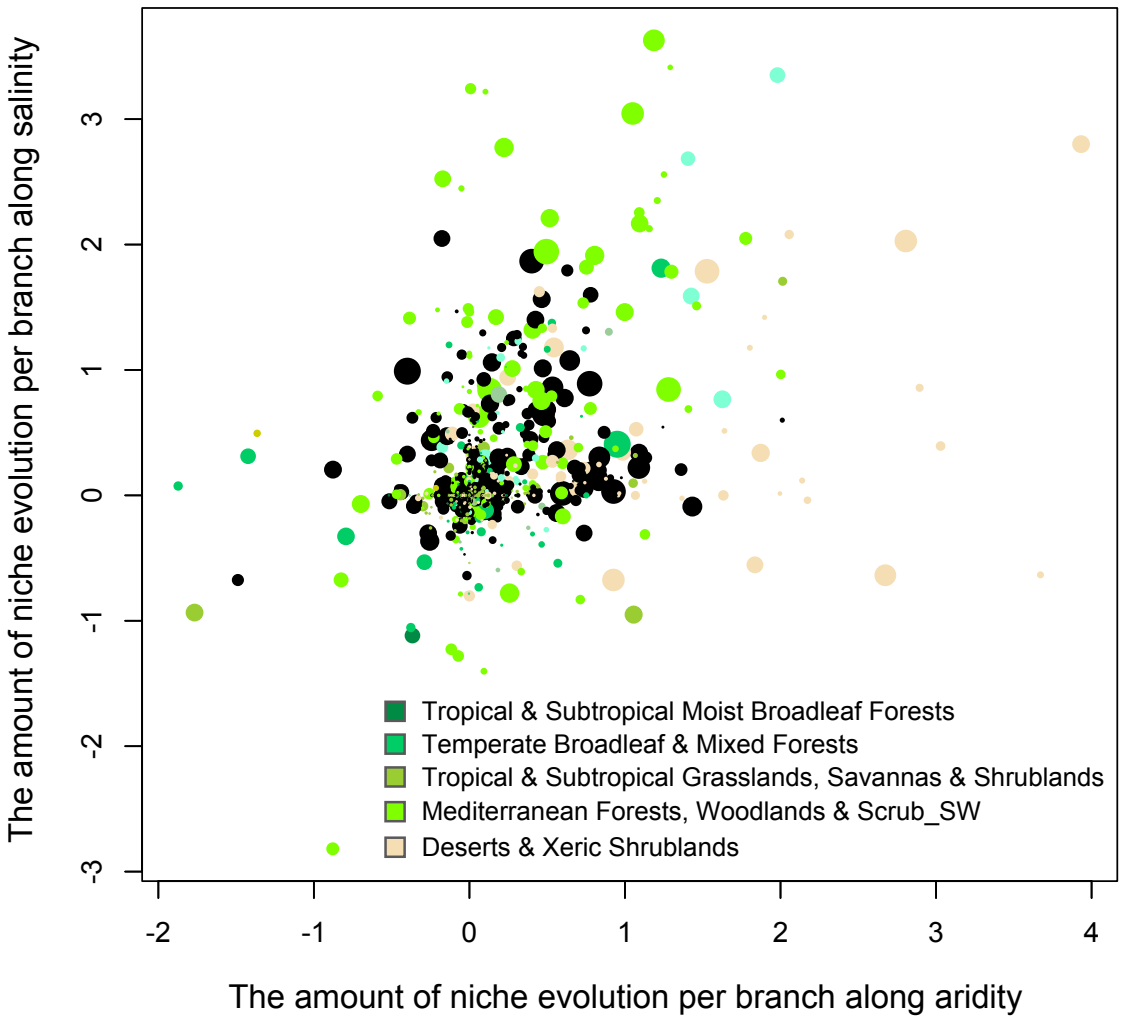



Figure 4. Correlation in the amount of niche evolution per branch along aridity and salinity. Each datapoint is a branch, with internal branches in black. The size of each datapoint is proportional to the tenth root of the branch length. Terminal branches are colored by the ecoregion that the majority of current presence location of the corresponding tip species fall in, using the same color scheme as the ecoregion map in Figure 1.