Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

INVERSION OF THE LATITUDINAL DIVERSITY GRADIENT AT HIGH TAXONOMIC LEVEL IN LIVERWORTS REVEALED BY A PHYLOGENETICALLY DECONSTRUCTIVE APPROACH

Jian Wang[1,2,3,*], Zun Dai[1], Thibault Kasprzyk[4], Xue Yao[1], Anders Hagborg[5], Lars Söderström[6], Jian Zhang[7], Alain Vanderpoorten[4] and Flavien Collart[4,8]

[1]*Bryology Laboratory, School of Life Science, East China Normal University, Shanghai 200241, China*
[2]*Shanghai Institute of Eco-Chongming (SIEC), 3663 Northern Zhongshan Road, Shanghai 200062, China*
[3]*Zhejiang Zhoushan Island Ecosystem Observation and Research Station, School of Ecological and Environmental Sciences, East China Normal University, Shanghai 200241, China*
[4]*University of Liège, Institute of Botany, B22 Sart Tilman, 4000 Liège, Belgium*
[5]*Department of Science and Education, The Field Museum, 1400 South Lake Shore Drive, Chicago, IL 60605-2496, USA*
[6]*Department of Biology, Norwegian University of Science and Technology, 7491 Trondheim, Norway*
[7]*School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, China, and*
[8]*Department of Ecology and Evolution (DEE), University of Lausanne, Lausanne, Switzerland*
**For correspondence. E-mail jwang@bio.ecnu.edu.cn*
#Jian Wang, Jian Zhang, Alain Vanderpoorten, and Flavien Collart contributed equally to this work.

BACKGROUND AND AIMS
Like numerous patterns in ecology and evolution, the latitudinal diversity gradient varies across phylogenetic levels. Yet, studies that investigate systematically how patterns and processes change at different phylogenetic levels, from the tips to the root, are still relatively scarce. Here, we test the hypothesis that, despite the high long-distance dispersal capacities of liverworts, which would be expected to result in the homogenization of their distributions, an increase of diversity with latitude persists at increasing phylogenetic level owing to macroclimatic niche conservatism since the earliest evolutionary history of the group.

METHODS
Liverwort distributions were scored for 450 operational geographical units worldwide. From the tips to the root, the phylogeny was sliced continuously to examine how taxonomic and phylogenetic diversity are correlated with latitude in a standardized way. Taxonomic diversity and mean phylogenetic distance among taxa were computed for each operational geographical unit at different phylogenetic levels and correlated with macroecological factors using spatial linear models.

KEY RESULTS
The correlation between taxonomic diversity and latitude shifted progressively from significantly negative at species level to non-significant, then significantly positive at the highest phylogenetic levels. Taxonomic diversity and mean phylogenetic distance were both significantly correlated with macroclimatic factors across all phylogenetic levels.

CONCLUSIONS
In contrast to the marked increase of angiosperm family diversity towards the tropics, the latitudinal diversity gradient evidenced at species level in liverworts decayed progressively at increasing phylogenetic level,

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

suggesting that phylogenetic niche conservatism has played a much weaker role in liverworts than in angiosperms. The inverted latitudinal diversity gradient towards the deepest phylogenetic levels lends support to the hypothesis that the earliest lineages diversified in extra-tropical conditions, explaining why, unlike in angiosperms, high species richness in the tropics is not associated with high phylogenetic diversity in liverworts. Our results highlight the extent to which a phylogenetically deconstructive approach allows for a better understanding of the accumulation of biodiversity through time.

KEY WORDS: Bryophyte, latitudinal diversity gradient, phylogenetic niche conservatism, phylogenetic scale, taxonomic diversity, phylogenetic diversity.

Introduction

The increase of species richness towards the tropics, a pattern known as the latitudinal diversity gradient (LDG), has long been considered to be one of most recurrent patterns in

ecology (Lomolino et al., 2010), with a remarkable consistency across space, habitats, taxonomic groups and time (Crane and Lidgard, 1989; Buckley and Jetz, 2007; Kreft and Jetz, 2007). Like numerous patterns in ecology and evolution, the LDG varies across phylogenetic levels (Graham et al., 2018).

In vertebrates, the diversity in higher-level taxa (e.g. families, orders) increases towards the tropics. At shallower taxonomic levels, variation in diversity with latitude varies from one taxon to another (Worm and Tittensor, 2018). Many young clades show negative richness–temperature slopes, with the ages of these clades coinciding with the expansion of temperate climate zones in the late Eocene (Buckley et al., 2010). The subsequent spread and diversification of these lineages to higher latitudes provided the possibility of different latitudinal distributions, resulting in a decay of the LDG towards finer phylogenetic scales (Worm and Tittensor, 2018).

In this context, a deconstructive macroecological approach, looking at how biodiversity patterns vary among areas and across taxonomic levels, allows for a better understanding of how biodiversity has accumulated through time, but also through space (Marquet et al., 2004). Analysing the relationship between the different biodiversity dimensions in a geographical context may, in fact, unveil the occurrence of particular ecological and evolutionary mechanisms in some geographical places or specific clades (Vásquez-Restrepo et al., 2023). In particular, the strength of the LDG typically varies longitudinally, being shallower at intermediate longitudes (-20° to 60°) and strongest at extreme longitudes (Kinlock et al., 2018). For example, tropical plants are more species rich in East Asia and the Americas compared with Africa as a result of differences in past and current climatic conditions and higher speciation rates in South America and Southeast Asia than in Africa (Couvreur, 2015; Hagen et al., 2021; de Miranda et al., 2022).

Despite the massive increase in phylogeny-based research over the last years, studies that systematically investigated how patterns and processes change with phylogenetic scale or use phylogenetic scale to examine the effects of aggregating or disaggregating taxa on global patterns are still relatively scarce (Worm and Tittensor, 2018; Chalmandrier et al., 2019; Vásquez-Restrepo et al, 2023). Here, we examine how two complementary diversity metrics, namely taxonomic diversity and phylogenetic diversity, vary along a phylogenetic time scale. Phylogenetic diversity measures the extent to which species in a community are closely or distantly related to each other. Integrating the shared evolutionary history of species, phylogenetic diversity can depart from taxonomic diversity. For instance, recent radiations are characterized by low phylogenetic but high taxonomic diversity, while rare long-distance dispersal events may contribute little to taxonomic diversity but substantially increase phylogenetic diversity (Davies and Buckley, 2011). In the context of the LDG, phylogenetic diversity specifically allows us to test one of the main driving forces of the pattern, namely tropical conservatism. The tropical conservatism hypothesis (Wiens and Donoghue, 2004) posits, based on the fact that tropical regions were more extensive until the mid-Tertiary, that temperate clades include lineages that diverged recently from tropical ones and share adaptations to cold, resulting in phylogenetically clustered temperate lineages (Kerkhoff et al., 2014).

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

We focus on liverworts, a group of land plants including 7271 species (Söderström et al., 2016) that, together with mosses and hornworts, compose the bryophytes. Although an increase of species richness away from the tropics was evidenced regionally (Rozzi et al., 2008; Mateo et al., 2016),

liverwort species richness peaks in tropical conditions at the worldwide scale (Wang et al., 2017). Liverworts are primarily wind dispersed by spores involved in chance transoceanic dispersal events, but the number of those events is largely lower than that of intracontinental speciation events, suggesting that the speciation rate is higher than the transoceanic dispersal rate (Patiño and Vanderpoorten, 2018). With increasing chances of a long-distance dispersal event during their evolutionary history, higher-level taxa are expected to exhibit larger distribution ranges. A decay of the LDG is therefore expected at increasing taxonomic depth, unless some lineages would have been restricted to specific environmental conditions. In particular, evidence for macroclimatic niche conservatism across liverworts floras worldwide (Collart et al., 2021) suggests that macroclimatic preferences are heritable, hampering tropical lineages from colonizing extra-tropical regions.

Specifically, we test the hypothesis that in liverworts, a latitudinal diversity gradient occurs at the species level and persists at increasing phylogenetic depth owing to macroclimatic niche conservatism, evidencing the role of the latter in shaping the observed patterns since the earliest evolutionary history of the group.

Materials and Methods

DATA SOURCES

We scored species distributions in operational geographical units (OGUs) from the most comprehensive database of liverwort species distributions available to date, which has been built in the context of the Early Land Plants Today project (Söderström et al., 2020). The distribution data are derived from a working database that, as of January 2019, included a bibliography of >11 000 references and >1 000 000 distribution records. Nomenclature follows Söderström et al. (2016). Based on this, we excluded 915 species, for which taxonomic evidence is conflicting (species with 'serious doubts', qualified with one star by Söderström et al., 2016) and scored the pres- ence/absence of 6331 accepted species names in 450 OGUs worldwide. Most of these units are defined by geopolitical boundaries, as commonly used in macroecological studies (e.g. Liu et al., 2023; Qian et al., 2024). They represent level-3 units (botanical countries) of the World Geographical Scheme for Recording Plant Distributions (https://www.tdwg.org/stand- ards/wgsrpd), with the largest countries (e.g. USA, Canada, China, Brazil and Australia) divided into states or provinces (level-4 units).

The data are available from https://doi.org/10.6084/ m9.figshare.22587199.v1.

Latitudinal variation in species richness was represented per band of 10°, including all the species recorded in the OGUs whose centroid was included in the band considered. Given that the proportion of land area decreases with increasing latitude in the Southern Hemisphere, we calculated species density for each latitudinal band as the number of species divided by the $\log_{10}$-transformed area (in kilometres squared) of the latitudinal band. To take the longitudinal variation of diversity patterns into account, we partitioned the data longitudinally into three

main regions: New World (North and South America), western Old World (Europe and Africa) and eastern Old World (Asia and Oceania) (Condamine et al., 2012).

We downloaded mean annual temperature (MAT) and mean annual precipitation (MAP) at 1-km resolution from CHELSA (https://chelsa-climate.org/bioclim). These data were complemented with information on the temporal variation of climate change, characterized by climate change velocity for MAT and MAP between Last Glacial Maximum (~21 000 years ago) and the present (1950–2000) at 2.5 arc-min resolution (Sandel et al., 2011). For each variable, we computed the mean for each of the 450 OGUs.

TAXONOMIC AND PHYLOGENETIC DIVERSITY

Taxonomic and phylogenetic diversity was computed for each OGU. Taxonomic diversity was computed by

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

counting, for each OGU, the number of species or higher-level taxa. For phylogenetic diversity, we computed the mean phylogenetic distance (MPD) among taxa within an OGU. MPD measures the mean phylogenetic distance separating all the taxa within an OGU from each other (i.e. a tree-wide assessment of relatedness among co-occurring taxa) and thus quantifies the overall relatedness of the taxa within an OGU. Unlike other metrics, such as phylogenetic diversity, MPD is independent from taxonomic diversity (Miller et al., 2017). An average value of the phylogenetic distance among taxa, MPD is a poor estimate of phylogenetic diversity when the number of taxa is low. In a global map of MPD among species across OGUs worldwide, some of the OGUs with the highest MPD are among the most species poor (Supplementary Data Fig. S1). This is particularly the case in some sub-Saharan African OGUs, which include only two, albeit highly divergent, species (one leafy and one thalloid species). Previous analyses showed that correlation patterns between phylogenetic diversity and environmental variation were similar when computed from datasets including all OGUs versus only OGUs with more than ten species, but with higher correlation coefficients with the reduced dataset (Collart et al., 2021). Therefore, MPD was not computed for 35 OGUs with a species richness of less than ten to avoid spurious estimates of MPD attributable to the low number of species included.

Although a phylogenomic analysis of 151 liverwort genera based on 228 nuclear genes was published most recently (Bechteler et al., 2023), we used the phylogeny of Laenen et al. (2014) based on the analysis of eight genes from all genomic compartments to compute MPD values. This phylogeny, whose topology is globally consistent with that of Bechteler et al. (2023), includes 303 genera, representing 84 % of the total extant generic diversity, i.e. the most densely sampled global liverwort phylogeny to date.

Given that the phylogeny includes a single species per genus, all congeneric species included in the distribution database were grafted onto the genus-level phylogeny. In the absence of well-resolved phylogenies at the species level, this approach has commonly been implemented in community phylogenetic analyses (for review, see Qian and Jin, 2021). Given that variation in the phylogenetic diversity of communities is driven mainly by ancestral branches (Hardy et al., 2012), phylogenetic metrics derived from species-level phylogenies resolved at the genus level with species being attached to their respective genera as polytomies are highly correlated with those derived from a phylogeny fully resolved at the species level (Qian and Jin, 2021). Here, we attempted to take the uncertainty of the diversity metrics obtained after the grafting of species to their respective genus by generating 100 trees with randomly resolved relationships among congeneric species using the 'add.species.to.genus' function of the phytools package (Revell, 2012). Although far from covering the entire range of alternative phylogenies, this approach allowed us, instead of focusing on a single arbitrarily selected phylogeny or attaching species to their respective genera in the phylogeny as polytomies, to examine the impact of the random grafting of species within their genera across 100 trees. To assess further the extent to which this approach could lead to a bias in the patterns of diversity observed, we contrasted them with those obtained from the analysis of the genus-level phylogeny of Laenen et al. (2014). Here, MPD was not computed for 294 OGUs that included fewer than ten genera, for reasons detailed above.

To determine whether species (or higher-level taxa) within an OGU are more or less phylogenetically related than expected by chance, characterizing phylogenetic clustering and overdispersion, respectively, 100 MPD values were generated from each of the 100 phylogenies (thus for a total of 10 000 MPD values) whose species were randomized among the tips to build the distribution of the null hypothesis, then compared with the observed MPD values.

STATISTICAL ANALYSES

We determined how species and higher-level taxa are distributed along latitudinal gradients to test the hypothesis that a LDG can be characterized at different nested taxonomic levels (i.e. whether there is an increase of the diversity of genera, families etc. towards the tropics). Instead of computing the numbers of higher-level taxa per OGU, to avoid the problem that taxonomic ranks are arbitrarily defined (Avise and Johns, 1999; Kraichak et al., 2017), we standardized the taxonomic level at which we characterize spatial diversity patterns by sampling taxa that are defined by a given age in the phylogeny. The regular levels at which the phylogeny was sliced do not, however, necessarily correspond to traditionally defined taxa, and we therefore refer hereafter to the phylogenetic level at which the analysis was conducted. To this end, we generated α-diversity-through-time profiles, analogous to the β-diversity-through-time framework (Groussin et al., 2017) at arbitrarily defined intervals of 25 Myr along the phylogenetic time scale. From the tips to the

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

root, the phylogenetic tree was sliced continuously, yielding, for each time period, clades that served as OTUs in downstream analysis. The phylogenetic tree was pruned to the desired depth by collapsing all descendent leaves of each of the branches encountered at 25 Myr periods. Accordingly, the branch lengths leading to these new terminal nodes progressively shortened and the number of OTUs progressively decreased towards the root (Supplementary Data Table S1). The geographical distribution of these branches was calculated as the union of the distributions of their descending leaves. For each 25 Myr period, we computed the correlation between taxonomic and phylogenetic diversity per OGU and the absolute value of the latitude of the centroid of each OGU.

To examine how taxonomic and phylogenetic diversity vary depending on macroclimatic variation, simultaneous autoregressive error (SAR) models, as implemented by the R package spatialreg (Bivand et al., 2023), were used to take the effects of spatial autocorrelation into account (Kissling and Carl, 2008; Zhang et al., 2013). We tested competing models involving all possible combinations of predictors. Competing models were compared on the basis of Akaike's information criterion (AIC) as proposed by Burnham and Anderson (2002), and the model with the lowest AIC was selected (Wagenmakers and Farrell, 2004). For the best-fitting model, we computed the relative importance of predictor variables by using the standardized partial regression coefficients for all predictor variables (Kissling and Carl, 2008; Zhang et al., 2013, 2016). We assessed the importance of individual variables to the model by summing the AIC weights across all subset models in which the variable was included (Burnham and Anderson, 2002). Then, variation in taxonomic and phylogenetic diversity across OGUs, considering its spatial component, was entered as a new response variable to perform SAR models again and obtain the standardized coefficient for each predictor (Belmaker and Jetz, 2011). Finally, we computed the P-value of the variables remaining in the model to identify 'confusing' variables, i.e. variables that would be kept in the model based on the AIC criterion, but would have a P-value > 0.05 (Sutherland et al., 2023). We summarized the standardized coefficients of variables in the best-fitting model and the weights of total variables together subsequently.

This analysis was repeated at increasing phylogenetic depth, i.e. at 25 Myr intervals along the phylogenetic tree. This approach does not account for the past distribution of climates when lineages arose, but simply seeks to determine whether current patterns of taxonomic and phylogenetic diversity at different phylogenetic levels vary depending on extant climatic conditions.

All statistical analyses were implemented in R v.4.2.3 (R Core Team, 2023), and the codes for all analyses implemented are available in the Appendix.

Results

Taxonomic diversity at the species level (i.e. species richness) of liverworts among the 450 OGUs ranged from 2 to 819, with an average of 14 species (Fig. 1). There was a clear increase of species density (i.e. species richness weighted by area) towards the tropics, as expressed by the clear increase of species density per latitudinal band of 10° to the tropics (Fig. 2). This global pattern resulted from a strong increase of species density towards the tropics in the Americas and Asia/Oceania, whereas in Europe/Africa, species density strikingly decreased between 15° and 35°, corresponding to the Sahel region. MPD globally tended to decrease from the poles to the tropics, shifting from higher to lower values than expected by chance (Supplementary Data Fig. S2), hence from significant phylogenetic overdispersion to clustering, except across Europe/ Africa (Figs 1B and 2B).

These patterns varied substantially at different phylogenetic levels (Fig. 3). The r^2 of the relationship between taxonomic richness per OGU and latitude decreased progressively at increasing phylogenetic level (Fig. 3C). The LDG globally collapsed from 50–75 Myr backwards, with most correlations between the number of lineages and latitude being insignificant across the 100 phylogenetic trees. About 100–125 Mya, the slope between taxonomic richness per OGU and latitude shifted from negative to positive (Fig. 3A). This shift was associated with a drastic increase of the r^2 in the Americas and Europe/Africa, but not in Asia/Oceania, to reach a value of 0.12 ($P < 0.001$) and 0.26 ($P < 0.001$) at 200 Mya, respectively (Fig. 3C). In contrast, the slope of the relationship between MPD and latitude remained positive across phylogenetic levels (Fig. 3B). The r^2 of this relationship tended to increase with phylogenetic depth in the Americas and Europe/Africa, but not in

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

Asia/ Oceania (Fig. 3D). The same patterns were recovered from the analysis using the genus-level phylogeny (Supplementary Data Fig. S3), evidencing the limited impact of the random grafting of species to their respective genus on the patterns of diversity observed.

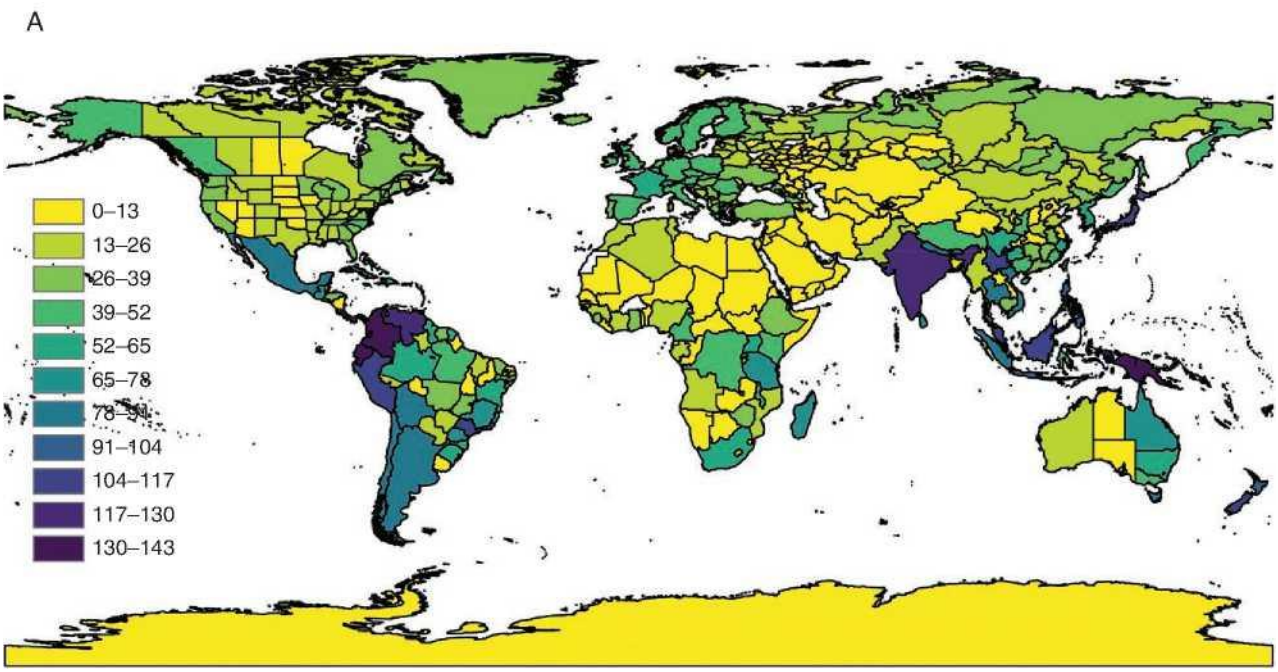
Significant models between taxonomic diversity, MPD and macroclimatic predictors were obtained across all phylogenetic depths, with r^2 ranging between 0.29 and 0.36 and between 0.09 and 0.27 for taxonomic diversity and MPD, respectively (Table 1). Competing models did not always exhibit significant differences (ΔAIC > 2), but yielded similar variable selections (Supplementary Data Table S2), with precipitation-related variables contributing globally more to models than temperature- related variables (Table 1).

Discussion

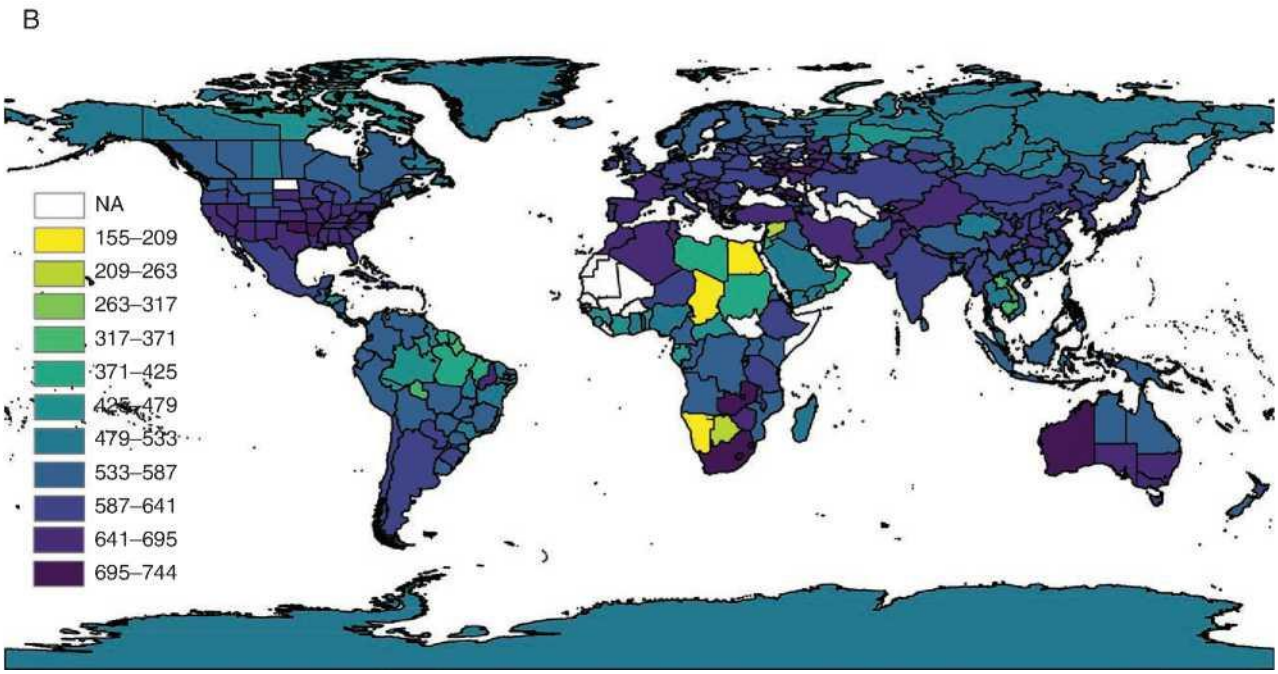
Our analyses reveal a sharp increase in liverwort species richness towards the tropics, evidencing the presence of a latitudinal diversity gradient in a group wherein it has long been debated owing to regional departures of this pattern (Rozzi et al., 2008; Mateo et al., 2016). The increase of species richness towards the tropics varied longitudinally, with a clear pattern across the Americas and Asia/Oceania, but not across Europe/ Africa. During the Last Glacial Maximum, the African rainforest area was reduced by ~84 %, i.e. to a much larger extent than its Amazonian counterpart, which probably shrank to 54 % of its present-day extent (Anhuf et al., 2006). Extinction processes are, furthermore, expected to be stronger in Africa than in Amazonia because of the smaller African rainforest area (~249 × 10^6 ha in Africa vs. 668 × 10^6 ha in Amazonia), resulting in smaller population sizes, which, in turn, increases extinction risk (Parmentier et al., 2007).

Such a clear latitudinal diversity pattern emerged despite the long-distance dispersal capacities of liverworts, which should have contributed to homogenize their distribution patterns. In fact, the significant correlation between phylogenetic diversity and macroclimatic variation that persisted through the entire phylogenetic time scale suggests, in line with previous evidence (Collart et al., 2021), that macroclimatic niche conservatism took place in the early evolutionary history of liverworts and continuously shaped their distribution patterns.

Fig. 1*. Map of the species richness (A) and phylogenetic diversity [B; measured as the mean phylogenetic distance (MPD) among species] across 450 OGUs in liverworts. Abbreviation: NA, OGUs for which MPD was not computed owing to low species richness (n < 10).*



Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

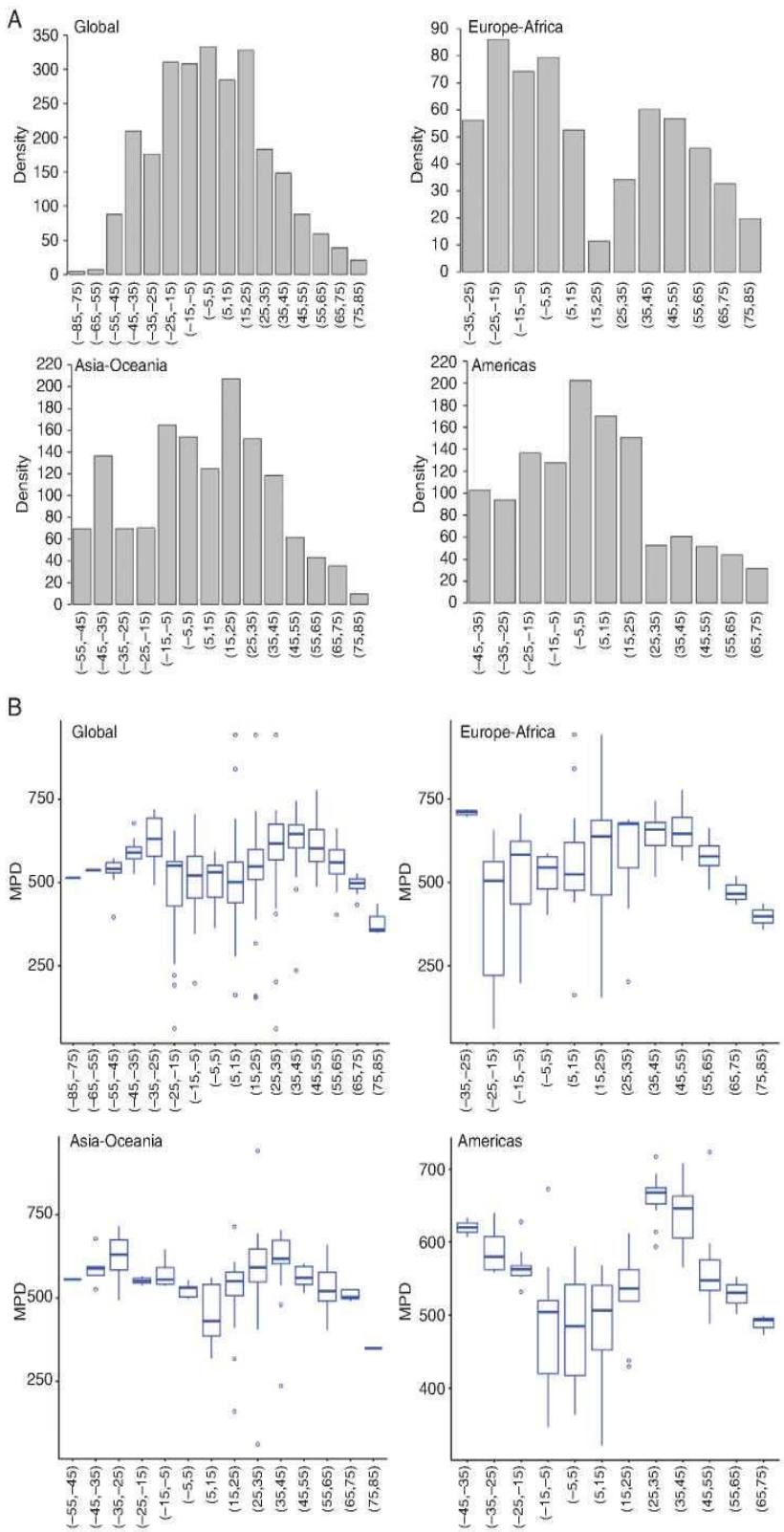


It is worth noting that significant correlations between liverwort phylogenetic diversity and macroclimatic conditions were revealed despite the fact that bryophytes are able to persist in microhabitats where a suitable microenvironment persists, long after the general climate of the region has changed (Schuster, 1983). As a result, bryophyte species distributions are typically disjunct (Patiño and Vanderpoorten, 2018). This suggests that, at a macroscale, such disjunctions associated with species microhabitat preferences do not prevent a global relationship between liverwort floras and macroclimatic conditions from emerging.

The latitudinal diversity gradient that emerged at the species level decayed progressively at increasing phylogenetic level, evidencing the wider range and weaker macroclimatic preference of high-level taxa. In fact, in our dataset, 41 % of liverwort species, but only 3 % of the genera and none of the families, exhibit a strictly tropical distribution. The erasure of the latitudinal diversity gradient at genus and family level in liverworts contrasts with the marked increase of angiosperm family diversity towards the tropics, shaped by strong tropical niche conservatism (Ramírez-Barahona et al., 2020). Although it has long been acknowledged that bryophytes exhibit higher long-distance dispersal capacities than spermatophytes (for review, see Patiño and Vanderpoorten, 2018), the rapid decay of the LDG in liverworts, together with the low, albeit significant, correlations between phylogenetic diversity and macroclimatic variation suggest that, by comparison with spermatophytes (Kerkhoff et al., 2014), liverwort distributions are little constrained by tropical niche conservatism. Reconstruction of the 'tropicality index', which reflects the proportion of species latitudinal range that falls within the tropics minus the proportion of the latitudinal range that falls outside of it (Kerkhoff et al., 2014), revealed, in fact, that many temperate lineages are phylogenetically clustered within tropical ones (Collart et al., 2021), evidencing multiple instances of successful transition from tropical environments to temperate ones. For example, all temperate species of the most species-rich liverwort family, Lejeuneaceae, are nested in clades comprising mainly tropical species (Wilson et al., 2007). Lejeunea, one of the largest liverwort genera, originated in the Neotropics and subsequently dispersed successfully into extra-tropical regions (Heinrichs et al., 2013; Lee et al., 2020), and the western European flora is composed largely of species of tropical origin (Patiño et al., 2015).

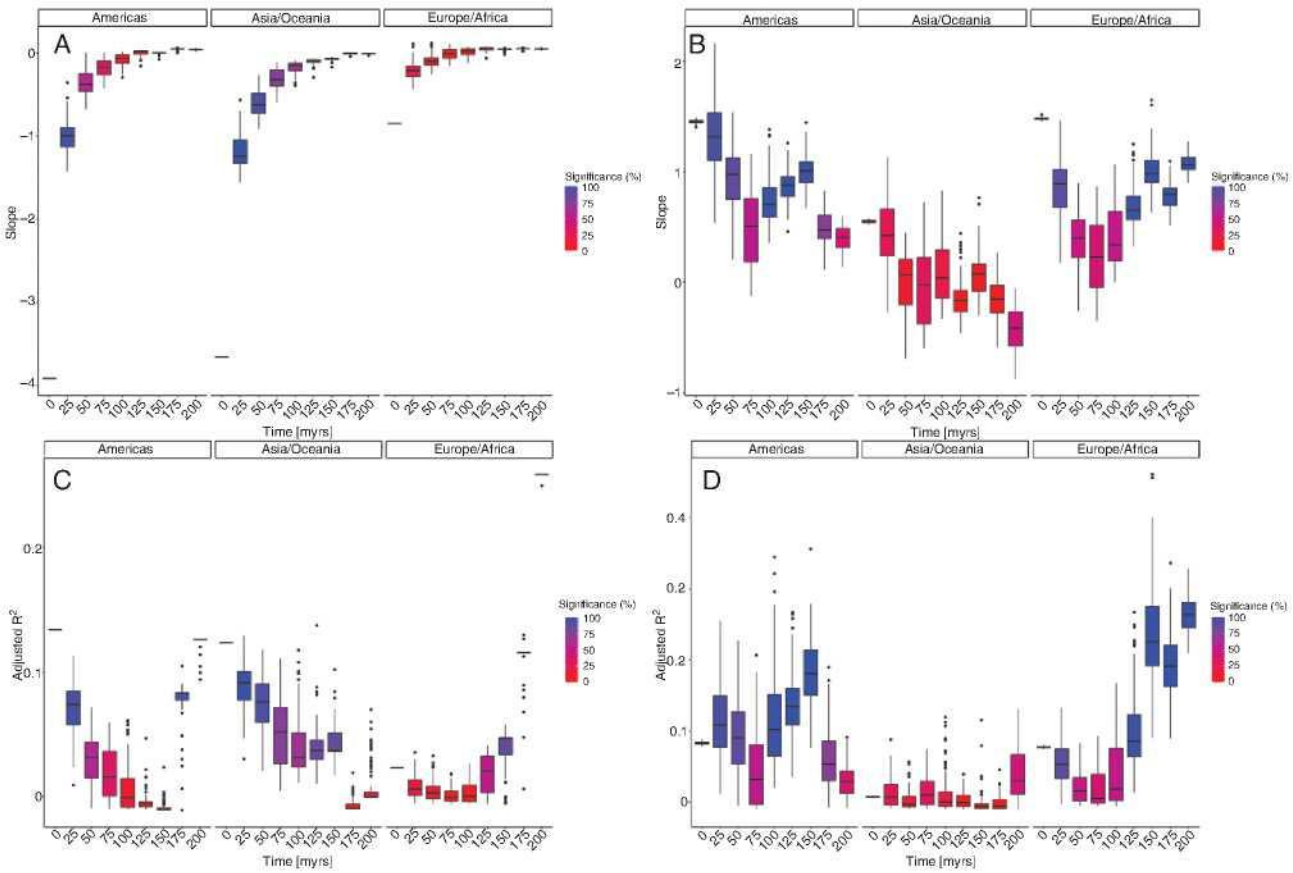
Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

***Fig. 2**. Latitudinal patterns of liverwort species density (species richness per latitudinal band divided by the log_{10}-transformed area; A) and mean phylogenetic distance (MPD) among species per OGU within each latitudinal band (B)*



Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

Fig. 3. *Variation of the slope (A, B) and r^2 (C, D) of the relationship between liverwort taxonomic richness (A, C) and MPD (B, D) per OGU and latitude (as an absolute value) through a phylogenetic time scale. Box plots [showing the first and third quartiles (upper and lower bounds), second quartile (centre), 1.5 × interquartile range (whiskers) and minima–maxima beyond the whiskers] show the variation across 100 trees, for which species relationships within each genus were randomized, and the colour code reflects the proportion of replicates for which the slope was significant.*



Precipitation-related variables contributed more to the observed correlation between taxonomic and phylogenetic diversity and macroclimatic factors than temperature-related variables, pointing to the strong water dependence of liverworts. Phylogenetic diversity decreased with precipitation- related variables. We tentatively suggest that drought exerts a strong evolutionary pressure on lineages, which developed convergent adaptations among unrelated lineages, thereby resulting in the observed phylogenetic overdispersion in dry areas. For example, hot, xeric areas typically host a combination of xerophilic, specialist complex thalloid liverworts (in particular, Ricciaceae) and generalist leafy liverworts, accounting for the comparatively high phylogenetic diversity, but low taxonomic diversity values in desertic areas of sub-Saharan Africa. In contrast, leafy liverworts prevail in extremely humid environments. The dominance of some hyperdiverse families, such as the Lejeuneaceae, in tropical rainforests leads to high levels of taxonomic diversity but low levels of phylogenetic diversity.

Most interestingly, the diversity of higher-level lineages towards the deepest phylogenetic levels (200 Myr and older) increased with latitude. This inverted LDG is reminiscent of the pattern observed today at the species level in temperate regions (southern South America, Rozzi et al., 2008; and Europe, Mateo et al., 2016) and, more globally, of the taxa that originated in cold conditions (Romdal et al., 2013).

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

TABLE 1. *Variation of the relationship between liverwort taxonomic richness and mean phylogenetic distance (a) and as a function of macroclimatic predictors (b) across 450 OGUs worldwide at increasing phylogenetic level. Time slice represents the phylogenetic level (in million years) at which the analysis was conducted.* b *and* W *are the standardized coefficient index (slope) and the summed Akaike information criterion (AIC) weights for each predictor across all possible spatial linear models, respectively. *P < 0.05, **P < 0.01 and ***P < 0.001. Abbreviations: MAP, mean annual precipitation; MAT, mean annual temperature; TV and PV, climate change velocity for mean annual temperature and annual precipitation, respectively, between the Last Glacial Maximum and the present.*

Time slice	0		25		50		75		100		125		150		175		200	
	b	W	b	W	b	W	b	W	b	W	b	W	b	W	b	W	b	W
(a)																		
MAT	–	0.38	−0.15*	0.81	−0.19**	0.95	−0.24***	0.99	−0.28***	1.00	−0.35***	1.00	−0.34***	1.00	−0.44***	1.00	−0.44***	1.00
MAP	0.40***	1.00	0.44***	1.00	0.43***	1.00	0.42***	1.00	0.41***	1.00	0.41***	1.00	0.42***	1.00	0.38***	1.00	0.37***	1.00
TV	–	0.48	–	0.33	–	0.29	–	0.30	–	0.35	–	0.42	0.12	0.49	–	0.40	0.17*	0.65
PV	−0.12*	0.64	−0.15**	0.86	−0.16**	0.92	−0.15**	0.93	−0.15**	0.92	−0.14**	0.92	−0.19**	0.92	–	0.56	−0.16*	0.73
r^2	0.29		0.31		0.31		0.30		0.32		0.34		0.36		0.36		0.35	
(b)																		
MAT	–	0.39	–	0.43	–	0.31	–	0.31	–	0.40	–	0.38	–	0.29	–	0.27	–	0.42
MAP	−0.12	0.70	−0.14*	0.58	−0.18**	1.00	−0.16**	0.96	−0.22***	1.00	−0.27***	1.00	−0.22***	0.99	−0.22***	0.99	−0.09	0.60
TV	0.20*	0.78	0.17*	0.64	0.13	0.52	0.14	0.63	0.12	0.55	0.11	0.57	0.17*	0.80	–	0.33	0.13	0.54
PV	0.17*	0.80	−0.17*	0.76	−0.17*	0.80	−0.19**	0.87	−0.18**	0.90	−0.24***	0.99	−0.23***	0.97	−0.12*	0.80	−0.16*	0.74
r^2	0.27		0.19		0.13		0.09		0.09		0.23		0.19		0.16		0.10	

Note: given that MAT was not included in the best-fitting model, there is no coefficient for it.

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

Although reconstructions of ancestral distribution ranges at deep nodes in the liverwort phylogeny are challenged by recurrent long-distance dispersal events (Laenen et al., 2018), it is tempting to see in the extra-tropical distributions of the centres of diversity of the basal lineages (e.g. Haplomitriaceae) an indication that such lineages originated in extra-tropical regions, where they would have become largely confined owing to macroclimatic niche conservatism. This is best exemplified by latitudinal gradients across South America, wherein a tropical-adapted 'modern' flora developed in the northern and central portions of the continent until at least mid-Tertiary, with few links to the flora of the southern part, which retained a large series of cool- and moisture-demanding Gondwanalandic elements (e.g. the families Haplomitriaceae and Blepharodophyllacaeae), largely shared with Australasia (Schuster, 1983). It is worth noting that although our analyses suggest that such taxa have been evolutionarily constrained by their macroclimatic niche, this does not necessarily mean that they persisted at locations where they are found today. Although this would be at odds with the high long-distance dispersal capacities of bryophytes, the distribution of continents, macroclimatic conditions and, most importantly, wind connectivity, which contributes more to bryophyte species distributions than geographical distance (Muñoz et al., 2004), varied substantially in geological time, impacting dispersal.

The distribution of such Gondwanalandic taxa in southern South America accounts for the inverted gradient of phylogenetic diversity across South America. Thus, high species richness in the tropics is not associated with high phylogenetic diversity. This was also evidenced in woody spermatophytes, wherein phylogenetically distant gymnosperm species contributed substantially to the high phylogenetic diversity of high-latitude communities, to which the bulk of gymnosperm species would be confined owing to phylogenetic niche conservatism (Massante et al., 2019).

The restriction of an inverse pattern of increasing MPD with latitude to the Americas raises the question of why the same pattern was not observed across Asia/Oceania, given that many Gondwanalandic taxa exhibit disjunct range disjunctions between southern South America and southern Australia and New Zealand (Schuster, 1983). In fact, comparatively high levels of MPD were observed in tropical Asia. This pattern corresponds to the occurrence of a series of isolated taxa of Gondwanalandic origin, such as the monogeneric, isolated family Delavayellaceae, which infused into the tropics via Indian plate migration and subsequently diversified following the elevation of the Himalaya (Schuster, 1983).

Conclusion

Liverworts exhibit striking variations of diversity through a phylogenetic time scale, highlighting how a deconstructive macroecological approach across phylogenetic levels allows for a better understanding of the accumulation of biodiversity through time. These changes highlight the fact that, even in organisms with high long-distance capacities, such as liverworts, changes in large-scale distribution patterns across phylogenetic levels, shaped by macroclimatic niche conservatism, are compatible with historical events associated with major diversification or extinction periods (Vanderpoorten et al., 2010). Most importantly, they suggest that although the latitudinal diversity gradient can vary among taxa (Cerezer et al., 2022), it can also vary substantially within the same taxon depending on the phylogenetic level investigated.

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

Supplementary Data

Supplementary data are available at Annals of Botany online and consist of the following.

Table S1: numbers of operational taxonomic units in the phylogenetic tree of liverworts sampled at genus level from Laenen et al. (2014) sliced at 25 Myr intervals. Table S2: model outputs of top three prediction models for mean phylogenetic distance (MPD) and species richness (SR). Abbreviations: AIC, Akaike's information criterion; Wr, Akaike weight. Figure S1: map of the phylogenetic diversity [measured as the mean phylogenetic distance (MPD) among species] across 450 operational geographical units in liverworts. Figure S2: phylogenetic structure, as expressed by mean phylogenetic distance (MPD), of liverwort communities per operational geographical unit, along a latitudinal gradient. Figure S3: variation of the slope and r^2 of the relationship between liverwort taxonomic richness (A) and MPD (B) per operational geographical unit and latitude (as an absolute value) through a phylogenetic time scale based on the original genus-level phylogeny of Laenen et al. (2014).

Funding

This study was funded by the National Natural Science Foundation of China (32070228), the Innovation Program of Shanghai Municipal Education Commission (2023ZKZD36) and the BEST (Biodiversity along Elevational gradients, Shifts and Transitions) research network (http://BEST-mountains.org).

Author Contributions

Jian Wang: Conceptualization (lead); Data curation (lead); Investigation (lead); Funding acquisition (lead); Methodology (supporting); Project administration (lead); Supervision (lead); Validation (lead); Visualization (lead); Writing – original draft (equal); Writing – review and editing (equal). Zun Dai: Formal analysis (lead); Methodology (supporting); Visualization (lead); Writing – original draft (supporting); Writing – review and editing (supporting). Thibault Kasprzyk: Formal analysis (supporting); Visualization (supporting). Xue Yao: Data curation (equal); Writing – review and editing (supporting). Anders Hagborg: Data curation (supporting). Lars Söderström: Data curation (supporting). Jian Zhang: Conceptualization (lead); Methodology (lead); Writing – original draft (equal); Writing – review and editing (supporting). Alain Vanderpoorten: Conceptualization (lead); Methodology (supporting); Writing – original draft (lead); Writing – review and editing (lead). Flavien Collart: Data curation (supporting); Formal analysis (equal); Visualization (equal).

Conflict of Interest

None declared.

Data Availability

Liverwort species distribution data used in this study are available from the Figshare repository at https://doi.org/10.6084/ m9.figshare.22587199. Molecular data sets and phylogenetic trees can be obtained from (http://purl.org/phylo/treebase/ phylows/study/TB2:S15950?x-access-code=98d94511d8c016 de8eae6f21c851826e&format=html). Climate data used in this study are available in the CHELSA climate database (https:// chelsa-climate.org/bioclim). Other original data included in the article and the Supplementary Data.

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

References

Anhuf D, Ledru MP, Behling H, *et al*. 2006. Paleo-environmental change in the Amazonian and African rainforest during the LGM. *Palaeogeography, Palaeoclimatology, Palaeoecology* 239: 510–527.

Avise JC, Johns GC. 1999. Proposal for a standardized temporal scheme of biological classification for extant species. *Proceedings of the National Academy of Sciences of the United States of America* 96: 7358–7363.

Bechteler J, Peñaloza-Bojacá G, Bell D, *et al*. 2023. Comprehensive phylogenomic time tree of bryophytes reveals deep relationships and uncovers gene incongruences in the last 500 million years of diversification. *American Journal of Botany* 110: e16249.

Belmaker J, Jetz W. 2011. Cross-scale variation in species richness–environment associations. *Global Ecology and Biogeography* 20: 464–474.

Bivand R, Piras G, Anselin L, *et al*. 2023. *Spatialreg, spatial models estimation and testing*. https://github.com/r-spatial/spatialreg (31 August 2023, date last accessed).

Buckley LB, Jetz W. 2007. Environmental and historical constraints on global patterns of amphibian richness. *Proceedings of the Royal Society B: Biological Sciences* 274: 1167–1173.

Buckley LB, Davies TJ, Ackerly DD, *et al*. 2010. Phylogeny, niche conservatism and the latitudinal diversity gradient in mammals. *Proceedings of the Royal Society B: Biological Sciences* 277: 2131–2138.

Burnham KP, Anderson DR. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*, 2nd edn. New York: Springer.

Cerezer FO, Machac A, Rangel TF, Dambros C S. 2022. Exceptions to the rule, relative roles of time, diversification rates and regional energy in shaping the inverse latitudinal diversity gradient. *Global Ecology and Biogeography* 31: 1794–1809.

Chalmandrier L, Pansu J, Zinger L, *et al*. 2019. Environmental and biotic drivers of soil microbial β-diversity across spatial and phylogenetic scales. *Ecography* 42: 2144–2156.

Collart F, Wang J, Patiño J, *et al*. 2021. Macroclimatic structuring of spatial phylogenetic turnover in liverworts. *Ecography* 44: 1474–1485.

Condamine L, Sperling FAH, Wahlberg N, *et al*. 2012. What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecology Letters* 15: 267–277.

Couvreur TLP. 2015. Odd man out, why are there fewer plant species in African rain forests? *Plant Systematics and Evolution* 301: 1299–1313.

Crane PR, Lidgard S. 1989. Angiosperm diversification and paleolatitudinal gradients in cretaceous floristic diversity. *Science* 246: 675–678.

Davies TJ, Buckley LB. 2011. Phylogenetic diversity as a window into the evolutionary and biogeographic histories of present-day richness gradients for mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366: 2414–2425.

de Miranda PLS, Dexter KG Swaine MD, *et al*. 2022. Dissecting the difference in tree species richness between Africa and South America. *Proceedings of the National Academy of Sciences of the United States of America* 119: e2112336119.

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

Graham CH, Storch D, Machac A. 2018. Phylogenetic scale in ecology and evolution. *Global Ecology and Biogeography* 27: 175–187.

Groussin M, Mazel F, Sanders JG, *et al*. 2017. Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. *Nature Communications* 8: 14319.

Hagen O, Skeels A, Onstein RE, Jetz W, Pellissier L. 2021. Earth history events shaped the evolution of uneven biodiversity across tropical moist forests. *Proceedings of the National Academy of Sciences of the United States of America* 118: e2026347118.

Hardy OJ, Couteron P, Munoz F, Ramesh BR, Pélissier R. 2012. Phylogenetic turnover in tropical tree communities: impact of environmental filtering, biogeography and mesoclimatic niche conservatism. *Global Ecology and Biogeography* 21: 1007–1016.

Heinrichs J, Dong S, Schäfer-Verwimp A, *et al*. 2013. Molecular phylogeny of the leafy liverwort *Lejeunea* (Porellales), evidence for a neotropical origin, uneven distribution of sexual systems and insufficient taxonomy. *PLoS One* 8: e82547.

Kerkhoff AJ, Moriarty PE, Weiser MD. 2014. The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 111: 8125–8130.

Kinlock NL, Prowant L, Herstoff EM, *et al*. 2018. Explaining global variation in the latitudinal diversity gradient, meta-analysis confirms known patterns and uncovers new ones. *Global Ecology and Biogeography* 27: 125–141.

Kissling W, Carl G. 2008. Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography* 17: 59–71.

Kraichak E, Crespo A, Divakar PK, Leavitt SD, Lumbsch HT. 2017. A temporal banding approach for consistent taxonomic ranking above the species level. *Scientific Reports* 7: 2297.

Kreft H, Jetz W. 2007. Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences of the United States of America* 104: 5925–5930.

Laenen B, Shaw B, Schneider H, *et al*. 2014. Extant diversity of bryophytes emerged from successive post-Mesozoic diversification bursts. *Nature Communications* 5: 5134.

Laenen B, Patiño J, Hagborg A, *et al*. 2018. Evolutionary origin of the latitudinal diversity gradient in liverworts. *Molecular Phylogenetics and Evolution* 127: 606–612.

Lee GE, Condamine FL, Bechteler J, *et al*. 2020. An ancient tropical origin, dispersals via land bridges and Miocene diversification explain the subcosmopolitan disjunctions of the liverwort genus *Lejeunea*. *Scientific Reports* 10: 14123.

Liu Y, Xu X, Dimitrov D, *et al*. 2023. An updated floristic map of the world. *Nature Communications* 14: 2990.

Lomolino MV, Riddle BR, Whittaker R, *et al*. 2010. *Biogeography*, 4th edn. Sunderland, MA: Sinauer Associates.

Marquet PA, Fernández M, Navarrete SA, *et al*. 2004. Diversity emerging, towards a deconstruction of biodiversity patterns. In: Heaney ML, Heaney MR. eds. *Frontiers of biogeography: new directions in the geography of nature*. Cambridge, UK: Cambridge Univerity Press, 192–209.

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

Massante JC, Götzenberger L, Takkis K, *et al*. 2019. Contrasting latitudinal patterns in phylogenetic diversity between woody and herbaceous communities. *Scientific Reports* 9: 6443.

Mateo RG, Broennimann O, Normand S, *et al*. 2016. The mossy north, an inverse latitudinal diversity gradient in European bryophytes. *Scientific Reports* 6: 25546.

Miller ET, Farine DR, Trisos CH. 2017. Phylogenetic community structure metrics and null models: a review with new methods and software. *Ecography* 40: 461–477.

Muñoz J, Felicisimo AM, Cabezas F, Burgaz AR, Martínez I. 2004. Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science* 304: 1144–1147.

Parmentier I, Malhi Y, Senterre B, Whittaker RJ. 2007. The odd man out? Might climate explain the lower tree alpha diversity of african rain forests relative to Amazonian rain forests? *Journal of Ecology* 95: 1058–1071.

Patiño J, Vanderpoorten A. 2018. Bryophyte biogeography. *Critical Reviews in Plant Sciences* 37: 175–209.

Patiño J, Carine M, Mardulyn P, *et al*. 2015. Approximate Bayesian computation reveals the crucial role of oceanic islands for the assembly of continental biodiversity. *Systematic Biology* 64: 579–589.

Qian H, Jin Y. 2021. Are phylogenies resolved at the genus level appropriate for studies on phylogenetic structure of species assemblages? *Plant Diversity* 43: 255–263.

Qian H, Qian S, Zhang J, Kessler M. 2024. Effects of climate and environmental heterogeneity on the phylogenetic structure of regional angiosperm floras worldwide. *Nature Communications* 15: 1079.

R Core Team. 2023. *R, a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.r-project. org (31 August 2023, date last accessed).

Ramírez-Barahona S, Sauquet H, Magallón S. 2020. The delayed and geographically heterogeneous diversification of flowering plant families. *Nature Ecology & Evolution* 4: 1232–1238.

Revell LJ. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.

Romdal TS, Araújo MB, Rahbek C. 2013. Life on a tropical planet, niche conservatism and the global diversity gradient. *Global Ecology and Biogeography* 22: 344–350.

Rozzi R, Armesto JJ, Goffinet B, *et al*. 2008. Changing biodiversity conservation lenses, insights from the sub-Antarctic non-vascular flora of southern South America. *Frontiers in Ecology and the Environment* 6: 131–137.

Sandel B, Arge L, Dalsgaard B, *et al*. 2011. The influence of late Quaternary climate-change velocity on species endemism. *Science* 334: 660–664.

Schuster RM. 1983. *New manual of bryology*. Nichinan: Hattori Botanical Laboratory.

Söderström L, Hagborg A, von Konrat M, *et al*. 2016. World checklist of hornworts and liverworts. *Phytokeys* 27: 1–828.

Söderström L, Hagborg A, von Konrat M. 2020. ELPT: early land plants today (ver. Jan 2019). In: Roskov Y, Kunze T, Orrell TM, *et al*. eds. *Species 2000 and ITIS Catalogue of Life, 2020-12-01*. Leiden, The Netherlands: Species 2000: Naturalis. www.catalogueoflife.org/data/ dataset/170394

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

Sutherland C, Hare D, Johnson PJ, Linden DW, Montgomery RA, Droge E. 2023. Practical advice on variable selection and reporting using Akaike information criterion. *Proceedings of the Royal Society B: Biological Sciences* 290: 20231261.

Vanderpoorten A, Gradstein SR, Carine MA, Devos N. 2010. The ghosts of Gondwana and Laurasia in modern liverwort distributions. *Biological Reviews of the Cambridge Philosophical Society* 85: 471–487.

Vásquez-Restrepo JD, Ochoa-Ochoa LM, Flores-Villela O, *et al*. 2023. Deconstructing the dimensions of alpha diversity in squamate reptiles Reptilia, Squamata across the Americas. *Global Ecology and Biogeography* 32: 250–266.

Wagenmakers EJ, Farrell S. 2004. AIC model selection using Akaike weights. *Psychonomic Bulletin & Review* 11: 192–196.

Wang J, Vanderpoorten A, Hagborg A, Goffinet B, Laenen B, Patiño J. 2017. Evidence for a latitudinal diversity gradient in liverworts and hornworts. *Journal of Biogeography* 44: 487–488.

Wiens JJ, Donoghue MJ. 2004. Historical biogeography, ecology, and species richness. *Trends in Ecology & Evolution* 19: 639–644.

Wilson R, Heinrichs J, Hentschel J, Gradstein SR, Schneider H. 2007. Steady diversification of derived liverworts under Tertiary climatic fluctuations. *Biology Letters* 3: 566–569.

Worm B, Tittensor DP. 2018. *A theory of global biodiversity*. Princeton: Princeton University Press.

Zhang J, Kissling WD, He F. 2013. Local forest structure, climate and human disturbance determine regional distribution of boreal bird species richness in Alberta, Canada. *Journal of Biogeography* 40: 1131–1142.

Zhang J, Hu J, Lian J, Fan Z, Ouyang X, Ye W. 2016. Seeing the forest from drones: testing the potential of lightweight drones as a tool for long-term forest monitoring. *Biological Conservation* 198: 60–69.

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

Appendix

CODES FOR ALL ANALYSES IMPLEMENTED IN THIS STUDY

(1) The code for generating Fig. 2B in this study. library(ggplot2)

```
dat = read.csv("MPD.csv",header = T)
Adat = dat[dat$continent == "Americas",]
    EAdat = dat[dat$continent == "EuropeAfrica",]
    AOdat = dat[dat$continent == "AsiaOceania",]
    dat$brks <- factor(dat$brks, levels=c("(-85,-75]", "(-65,55]", "(-55,-45]", "(-45,-35]", "(-35,-25]", "(-25,-15]", "(-15,-5]","(-5,5]", "(5,15]", "(15,25]", "(25,35]", "(35,45]", "(45,55]", "(55,65]", "(65,75]", "(75,85]"))
    ggplot(dat, aes(x = brks, y = MPD, group = brks_ mid)) + geom_boxplot() + theme_minimal() +
    theme(
    plot.background = element_rect(fill = "white"),
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    panel.border = element_rect(color = "white", fill = NA),
    axis.line = element_line(color = "black"),
    axis.ticks = element_line(color = "black"),
    axis.text.x = element_text(angle = 90,vjust = 0.5, hjust = 1)
    )
    EAdat$brks <- factor(EAdat$brks, levels=c("(-35,-25]", "(-25,-15]", "(-15,-5]", "(-5,5]", "(5,15]", "(15,25]", "(25,35]", "(35,45]", "(45,55]", "(55,65]", "(65,75]", "(75,85]"))
    ggplot(EAdat, aes(x = brks, y = MPD, group = brks_ mid)) + geom_boxplot() + theme_minimal() +
    theme(
    plot.background = element_rect(fill = "white"),
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    panel.border = element_rect(color = "white", fill = NA),
    axis.line = element_line(color = "black"),
    axis.ticks = element_line(color = "black"),
axis.text.x = element_text(angle = 90,vjust = 0.5, hjust = 1) )
    AOdat$brks <- factor(AOdat$brks, levels=c("(-55,-45]", "(-45,-35]", "(-35,-25]", "(-25,-15]", "(-15,-5]", "(-5,5]", "(5,15]", "(15,25]", "(25,35]", "(35,45]", "(45,55]", "(55,65]", "(65,75]", "(75,85]"))
    ggplot(AOdat, aes(x = brks, y = MPD, group = brks_ mid)) + geom_boxplot() + theme_minimal() +
    theme(
    plot.background = element_rect(fill = "white"),
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    panel.border = element_rect(color = "white", fill = NA),
    axis.line = element_line(color = "black"),
    axis.ticks = element_line(color = "black"),
    axis.text.x = element_text(angle = 90,vjust = 0.5, hjust = 1)
    )
    Adat$brks <- factor(Adat$brks, levels=c("(-45,-35]", "(-35,-25]", "(-25,-15]", "(-15,-5]", "(-5,5]", "(5,15]", "(15,25]", "(25,35]", "(35,45]", "(45,55]", "(55,65]", "(65,75]"))
    ggplot(Adat, aes(x = brks, y = MPD, group = brks_ mid)) + geom_boxplot() + theme_minimal() +
    theme(
    plot.background = element_rect(fill = "white"),
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    panel.border = element_rect(color = "white", fill = NA),
    axis.line = element_line(color = "black"),
    axis.ticks = element_line(color = "black"),
axis.text.x = element_text(angle = 90,vjust = 0.5, hjust = 1) )
```

(2) The code for generating Fig. 3 in this study.

```
################################################
##################
    # Wang et al, adapted from Collart et al, 2021. Ecography #
    # MPD through time #
    ################################################ ##################
    # setwd("PutYourWdHere")
    # ## library ###
    library(ape)
    library(phytools)
    library(picante)
    library(vegan)
    library(betapart)
    library(adespatial)
    library(ade4)
    library(spdep)
    library(ggplot2)
    library(reshape2)
    library(lme4)
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
library(lmerTest)
library(rsq)
library(colorRamps)
########################################
# ## Functions used for the computation ###
########################################
# ## Function 1 ###
# Takes the node path for each species
noeudListe <-function(i,liste){
unlist(lapply(liste,function(x) × [i]))
}
# ## Function 2 ###
#     Checks if the node is present in the node path for a species
present<- function(k,data,i){
sum(data[[k]]==i)
}
# ## Function 3 ###
# Gives the position of present species in a region
estEgaleA1 <-function(vec){
data.frame(which(vec==1))
}
# ## Function 4 ###
# Takes All the species name present for a same node
nomASuppr <- function(liste,nomCol){
nomCol[liste[,1]]
}
##################
############################
# ## Replicat computation ###
############################
for(z in 1:100){# 100 repetitions
################################################ ######
# ## Data loading and preparation for the computation #
################################################ ######
# # Adapted From Collart et al, 2021. Ecography
print(paste0("start of replicate ",z))
treeLiver = read.tree("data/changed_phylo_liverwort_la- grange.tree") # Phylogenetic tree at
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

genus level

dataLiver = read.csv("data/Liverworts-database-to-analyse .csv",row.names = 1) # Species matrix (columns = OGUs and rows = species)

dataLiver <- dataLiver[,which(apply(dataLiver,2,sum)>=2)]

Restrained the matrix to minimum 2 sp per OGUs but later filtering at 10

Remove unused OGUs due to their old names (e.g. Yugoslavia) or the lack of environmental data

aRetirer <- c("X13YUG","X40WHM","X14RUS","X40ASS","X85MST","X85PAT","X85TDF","X90MPE","X21SEL", "X28STH","X11CZL","X30EVE","X31MGW","X33NCS"," X60GIL","X81SWC","X84BTR")

for(i in 1: length(aRetirer)){

dataLiver <- dataLiver[,colnames(dataLiver)!=aRetirer[i]]

}

Remove species that is not present in any OGUs

dataLiver <- dataLiver[which(apply(dataLiver,1,sum)!=0),]

speciesLiver = row.names(dataLiver) #Species name

commLiver<-t(dataLiver) #transposition of the dataset

phyloLiver<- force.ultrametric(treeLiver)

Remove duplicates that form monophyletic groups

dropping_before <- read.table("data/dropping_before.txt", stringsAsFactors = FALSE, header = F)$V1 # Vector containing the species names that are the same

dropping_after <- read.table("data/dropping_after.txt", stringsAsFactors = FALSE, header = F)$V1 # Vector containing the species name used for realising the phylogenetic tree

phyloLiver <- drop.tip(phyloLiver, dropping_before)

Randomly add species to genus

for(i in 1:length(speciesLiver))

{

phyloLiver<- force.ultrametric(phyloLiver)

phyloLiver<-add.species.to.genus(phyloLiver,speciesLiver [i],where="random")

}

#Remove duplicates

phyloLiver <- drop.tip(phyloLiver, dropping_after)

write.tree(phyloLiver,paste0(z,"/outputs/liverwortTree.tre"))

Splitting the OGUs into different regions

tabGroupeLiver <- read.csv("data/tableau-centro-trop2. csv") # Contains OGU center coordinates with the delimitation T/ETN/ETS

aSupp <-c() # vector that will contain position of OGUs that is not used in the centroid dataset

Give position of unused OGUs

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
for(i in 1: length(as.character(tabGroupeLiver[,3]))){
if(sum(colnames(dataLiver) == paste0("X",as.character(tab
GroupeLiver[i,3]))) == 0){
aSupp <- c(aSupp,i)
}
}
# Remove it from the table
tabGroupeLiver<-tabGroupeLiver[-aSupp,]
rownames(tabGroupeLiver) = tabGroupeLiver$a.garder
# # Adaptation for computation through time
phyloLiver$node.label = c(1:phyloLiver$Nnode) + length(p hyloLiver$tip.label) #add labels to node matching their number in object phylo
nodeagesLiver <- branching.times(phyloLiver) #min age of branch starting up to node
# add abundances up the phylogeny #
commwnodesLiver<-as.data.frame(commLiver)
commwnodesLiver[,length(commLi ver[1,]) + (1:phyloLiver$Nnode)] = 0
colnames(commwnodesLiver[,(length(co mmLiver[1,]) + (1:phyloLiver$Nnode))])<- length(commLiver[1,]) + c(1:phyloLiver$Nnode)
# identify node paths from root to each tip #
npLiver<-nodepath(phyloLiver)
# compute abundance for each node #
for(i in 1:length(commLiver[,1])){
for(t in 1:length(phyloLiver$tip.
label)) if(commwnodesLiver[i,t] > 0)
commwnodesLiver[i,npLiver[[t]][1:(length(npLiver[[t]])-1)]] = commwnodesLiver[i,npLiver[[t]][1:(length(npLiver[[t]])-1)] ] + commwnodesLiver[i,npLiver[[t]][length(npLiver[[t]])]]
}
write.table(commwnodesLiver,paste0("old_data/nouveau/ commwnodes_",z,".txt"),sep="\t")
commwnodesLiver <- read.table(paste0("old_data/nouveau/ commwnodes_",z,".txt"),sep="\t")
# compute inf/sup age of each branch #
agenLiver = as.data.frame(matrix(ncol = 2,nrow = length(co mmwnodesLiver[1,])))
colnames(agenLiver)=c("sup","inf")
edgeLiver<-rbind(phyloLiver$edge,c((length(phyloLiver$ tip.label) + 1),(length(phyloLiver$tip.label) + 1))) #make an edge table including the root
edgeLiver<-edgeLiver[order(edgeLiver[,2]),] #ordered
phylo$edge by descendent tip/node number
agenLiver$inf[1:length(commLiver[1,])] = 0 #inf age of tips = 0
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
agenLiver$sup[] = nodeagesLiver[edgeLiver[,1]- length(phyloLiver$tip.label)]
agenLiver$inf[(length(phyloLiver$tip.label) + 1):length(age nLiver$inf)] = nodeagesLiver
write.table(agenLiver,paste0("old_data/nouveau/age_ nodes_inf_sup_",z,".txt"),sep="\t")
#write this branch age table
agenLiver <- read.table(paste0("old_data/nouveau/age_ nodes_inf_sup_",z,".txt"),sep="\t")
# identify node paths from root to each tip #
npLiver<-nodepath(phyloLiver)
# Dataset preparation for the computation through time
dataPlot <- as.data.frame(matrix(0,ncol = length(phyloLiver
$tip.label) + 3,nr = nrow(agenLiver)))
dataPlot[1] = 1:nrow(agenLiver) # name of the node
colnames(dataPlot)[1]="NodeAge"
colnames(dataPlot)[4:(length(phyloLiver$tip.
label) + 3)] = phyloLiver$tip.label
colnames(dataPlot)[2:3]=c("ageSup","ageInf")
dataPlot[,2:3] = agenLiver #Age of the node
# species present for each node #
for(i in 1:nrow(dataPlot)){
spNode <- phyloLiver$tip.label[sapply(1:length(phyloLiver
$tip.label),present,data = npLiver,i = i)==1]
dataPlot[i,spNode] = 1
}
maxi <- max(unlist(lapply(npLiver,length)))
nomSp <- colnames(dataPlot[,-c(1:3)])
PhyloTime <- phyloLiver
dataTime <- commLiver
# # Create a matrix where node is present per OGU
numberNode <- nrow(phyloLiver$edge)
for(i in 1:(numberNode + 1)){
noeud = i
noeudPresent<-lapply(lapply(npLiver,function(x,noeud) {if(length(which(x == noeud))!=0){return(x)}},noeud = noeu d),length)
nomSpRegroup <- phyloLiver$tip.
label[which(noeudPresent > 0)]
petitTab <- commLiver[,nomSpRegroup]
if(is.matrix(petitTab)){
colARajouter<- apply(petitTab,1,sum)
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
colARajouter[colARajouter > 1] = 1
}
else{colARajouter <- petitTab}
dataTime <- data.frame(dataTime,colARajouter)
colnames(dataTime)[colnames(dataTime)==“colARajou ter”] = noeud
}
dataTime <- dataTime[,-c(1:2346)] # Remove species only keep node name
colnames(dataTime)=sub(pattern = “X”,replacement = “”,colnames(dataTime))
write.table(dataTime,paste0(“old_data/nouveau/ commLiveramongNodes_”,z,“.txt”),sep=“\t”)
dataTime <- read.table(paste0(“old_data/nouveau/
commLiveramongNodes_”,z,“.txt”),sep=“\t”)
colnames(dataTime) = sub(“X”,“”,colnames(dataTime))
rm(nodeagesLiver,commwnodesLiver,npLiver,maxi,spN ode)
################################################ ######
####################
# # Temporal loop ##
####################
resuAlphaTaxo <- NULL
resuAlphaPhylo <- NULL
PhyloTime<-phyloLiver
for(i in 1:201){
age = i-1 #began at time = 0 to 200 Mya
if(i == 1){age = 0.0000000001}
# Node Selection based on time
commLiverTime<-dataPlot[dataPlot$ageInf < age & dataPlot$ageSup>=age,-c(1:3)]
# regroup all the species that are in the same node to consider it as only one “species”
nombreSpAregrouper <- which(apply(commLiverTime,1, sum)>=2)
dataSP <- apply(commLiverTime,1,estEgaleA1)
spName <- do.call(rbind,lapply(dataSP,nomASuppr,nom Col = nomSp))
spNameToChange <- spName[,1]
dataTime2 <-dataTime[,rownames(commLiverTime)] # Only keeps species that will be used for
the computation
colnames(dataTime2) = spNameToChange
if(length(nombreSpAregrouper)!=0){#Regroup some species When we go through time
for(j in 1: length(rownames(commLiverTime))){
node <- which(phyloLiver$edge[,1]==as.
numeric(rownames(commLiverTime))[j])
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
if(length(node)!=0){
PhyloTime$edge.length[node] = 0 #decrease the previous branch to 0
}
}
}
b <- table(spNameToChange)
ToMerge <- names(b)[b > 1]
if(length(ToMerge) > 0){
for(g in 1:length(ToMerge)){
SeveralCol <- apply(dataTime2[,grep(ToMerge[g],colnames (dataTime2))],1,sum)
dataTime2 <- dataTime2[,-grep(ToMerge[g],colnames(data
Time2))]
dataTime2 <- cbind.data.frame(dataTime2,SeveralCol)
colnames(dataTime2)[ncol(dataTime2)] = ToMerge[g]
}
}
dataTime2<- dataTime2[which(apply(dataTime2,1,sum) > 1),]       # Keeps
OGUs that have at least 1 "species"
dataTime2<-dataTime2[,which(apply(dataTime2,2,
sum) > 0)] # Keeps OGUs that have at least 1 "species"
if(i==1){taillechgt <-c(i,nrow(dataTime2))}
if(i!=1){taillechgt <-rbind(taillechgt,c(i,nrow(dataTime2)))}
if(i %in% seq(1201,by = 25)){#Each 25 Myrs
print(age)
# Remove unused species
PhyloTime2 <- keep.tip(PhyloTime,colnames(dataTime2))
tabledivLiverTime2 <- specnumber(dataTime2)
write.table(tabledivLiverTime2,paste0("outputs/
AlphaTaxo_",z,"_time",(i-1),".txt"),sep="\t")
lati <- tabGroupeLiver[sub("X","",names(tabledivLiverT ime2)),]
lati <- lati$ycoord
mod <- lm(tabledivLiverTime2~abs(lati))
g <- summary(mod)
interAlphaTaxo <- cbind.data.frame(age=(i-1),adj.r.squared = g$adj.r.squared,slope =
g$coefficients["abs(lati)","Estim ate"], pval = g$coefficients["abs(lati)","Pr(>|t|)"])
resuAlphaTaxo <- rbind.data.frame(resuAlphaTaxo,interAlp haTaxo)
dis <- cophenetic(PhyloTime2)
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
alphaPhylo <- mpd(dataTime2,dis)
names(alphaPhylo) = names(tabledivLiverTime2)
write.table(alphaPhylo,paste0("outputs/AlphaPhylo_",z,"_
time",(i-1),".txt"),sep="\t")
mod <- lm(alphaPhylo~abs(lati))
g <- summary(mod)
interAlphaPhylo <- cbind.data.frame(age=(i-1),adj.r.squared = g$adj.r.squared,slope = g$coefficients["abs(lati)","Estim ate"], pval = g$coefficients["abs(lati)","Pr(>|t|)"])
resuAlphaPhylo <- rbind.data.frame(resuAlphaPhylo,interA lphaPhylo)
}
}
write.table(resuAlphaPhylo,paste0("outputs/LatiCorr_ AlphaPhylo_",z,".txt"),sep="\t")
write.table(resuAlphaTaxo,paste0("outputs/LatiCorr_
AlphaTaxo",z,".txt"),sep="\t")
print(paste0("end of the replicat ",z))
}
############################
# ## Output analyses ###
############################
# # Alpha taxonomy through time LMM test
filesAlpha <- list.files(pattern = "AlphaTaxo_")
fin <-list()
for(i in 1:length(filesAlpha)){
Age <- sub(".*time","",filesAlpha[i])
Age <- sub(".txt","",Age,fixed = T)
repet <- sub("_time.*","",filesAlpha[i])
repet <- sub("AlphaTaxo_","",repet,fixed = T)
inter <- cbind.data.frame(read.
table(filesAlpha[i],sep="\t"),Age,repet)
inter$OGU = row.names(inter)
row.names(inter) = NULL
fin[[i]] <- inter
}
fin <- do.call(rbind.data.frame,fin)
fin$Age = as.numeric(fin$Age)
test <-fin
test <- test[test$Age=="0" & test$repet=="1",c(1,4)]
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
colnames(test)[1] =“AlphaTaxo”
write.table(test,“AlphaTaxo_OGU.txt”,sep=“\t”)
# # Alpha Phylo through time LMM test
filesAlpha <- list.files(pattern = “AlphaPhylo_”)
filesAlpha <- filesAlpha[-grep(“LatiCorr”,filesAlpha)]
fin <-list()
for(i in 1:length(filesAlpha)){
Age <- sub(“.*time”,“”,filesAlpha[i])
Age <- sub(“.txt”,“”,Age,fixed = T)
repet <- sub(“_time.*”,“”,filesAlpha[i])
repet <- sub(“AlphaPhylo_”,“”,repet,fixed = T)
inter <- cbind.data.frame(read.
table(filesAlpha[i],sep=“\t”),Age,repet)
inter$OGU = row.names(inter)
row.names(inter) = NULL
fin[[i]] <- inter
}
fin <- do.call(rbind.data.frame,fin)
fin$Age = as.numeric(fin$Age)
test <-fin
test <- test[test$Age==“0” & test$repet==“1”,c(1,4)]
colnames(test)[1] =“mpd”
write.table(test,“mpd_OGU.txt”,sep=“\t”)
###
# # Data preparation for Fig. 1
library(foreign)
plot <- read.dbf(“Map/liverworts-true-final.dbf”) # Data
from Collart et al, 2021
MPD <- read.table(“mpd_OGU.txt”)
Taxo <- read.table(“AlphaTaxo_OGU.txt”,sep=“\t”)
plot$MPD = NA
plot$Taxo = NA
for(i in 1:450){#For each OGU, we store the MPD and species richness
plot$MPD[plot$A_GARDER == sub(“X”,“”,MPD$OGU[i]) &!is.na(plot$A_GARDER)] = MPD$mpd[i]
plot$Taxo[plot$A_GARDER == sub(“X”,“”,Taxo$OGU[i])
&!is.na(plot$A_GARDER)] = Taxo$AlphaTaxo[i]
}
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
write.dbf(plot,"Map/liverworts-true-final.dbf")
#################
# # MPD~Latitude for 3 longitudinal band
##############################
## Filtering n = 10 min
tabGroupeLiver <- read.csv("data/tableau-centro-trop2. csv") # Contains OGU center coordinates with the delimitation T/ETN/ETS
tabGroupeLiver$LongiBand = "0"
tabGroupeLiver$LongiBand[tabGroupeLiver$xco
ord < (-60)] = "Americas" ## Amear
tabGroupeLiver$LongiBand[tabGroupeLiver$xcoord > 60]
= "Asia/Oceania" ## Amear
tabGroupeLiver$LongiBand[tabGroupeLiver$xcoord <= 60 & tabGroupeLiver$xcoord >= (-60)] = "Europe/Africa" ## Amear
tabGroupeLiver$a.garder = paste0("X",tabGroupeLiver$a. garder)
filesAlphaTaxo <- list.files(pattern = "AlphaTaxo_",path="o utputs",full.names = T)
filesAlphaPhylo <- list.files(pattern = "AlphaPhylo_",path = "outputs",full.names = T)
filesAlphaPhylo <- filesAlphaPhylo[-grep("LatiCorr",filesA lphaPhylo)]
resuAlphaPhylo <- NULL
resuAlphaTaxo <- NULL
resuAlphaPhylo <- NULL
finTaxo <-list()
finPhylo <- list()
for(i in 1:length(filesAlphaTaxo)){
Age <- sub(".*time","",filesAlphaTaxo[i])
Age <- sub(".txt","",Age,fixed = T)
repet <- sub("_time.*","",filesAlphaTaxo[i])
repet <- sub("outputs/AlphaTaxo_","",repet,fixed = T)
inter <- cbind.data.frame(read.
table(filesAlphaTaxo[i],sep="\t"),Age,repet)
colnames(inter)[1]="AlphaTaxo"
inter$OGU = row.names(inter)
row.names(inter) = NULL
ToKeep <- inter$OGU[inter$AlphaTaxo>=10]
inter <- inter[inter$OGU %in% ToKeep,]
finTaxo[[i]] = inter
tab2 <- tabGroupeLiver[tabGroupeLiver$a.garder %in% inter$OGU,]
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
test <- merge(inter,tab2,by.x="OGU",by.y= "a.garder")
test$LongiBand = as.factor(test$LongiBand)
#Perform a linear model for each Longitudinal band
mod <- lm(AlphaTaxo~abs(ycoord),data = test)
g <- summary(mod)
interAlphaTaxo <- cbind.data.frame(age = Age,LongiBand = "FULL",repet = repet,
adj.r.squared = g$adj.r.squared,
slope = g$coefficients["abs(ycoord)","Estimate"],
pval = g$coefficients["abs(ycoord)","Pr(>|t|)"])
resuAlphaTaxo <- rbind.data.frame(resuAlphaTaxo,interAlp haTaxo)
for(j in levels(test$LongiBand)){
petMat <- test[test$LongiBand == j,]
mod <- lm(AlphaTaxo~abs(ycoord),data = petMat)
g <- summary(mod)
interAlphaTaxo <- cbind.data.frame(age = Age,LongiBand = j,repet = repet,
adj.r.squared = g$adj.r.squared,
slope = g$coefficients["abs(ycoord)","Estimate"],
pval = g$coefficients["abs(ycoord)","Pr(>|t|)"])
resuAlphaTaxo <- rbind.data.frame(resuAlphaTaxo,interAlp haTaxo)
}
Age <- sub(".*time","",filesAlphaPhylo[i])
Age <- sub(".txt","",Age,fixed = T)
repet <- sub("_time.*","",filesAlphaPhylo[i])
repet <- sub("outputs/AlphaPhylo_","",repet,fixed = T)
inter <- cbind.data.frame(read.
table(filesAlphaPhylo[i],sep="\t"),Age,repet)
colnames(inter)[1]="mpd"
inter$OGU = row.names(inter)
row.names(inter) = NULL
inter <- inter[inter$OGU %in% ToKeep,]
finPhylo[[i]] = inter
tab2 <- tabGroupeLiver[tabGroupeLiver$a.garder %in% inter$OGU,]
test <- merge(inter,tab2,by.x="OGU",by.y= "a.garder")
test$LongiBand = as.factor(test$LongiBand)
mod <- lm(mpd~abs(ycoord),data = test)
g <- summary(mod)
interAlphaPhylo <- cbind.data.frame(age = Age,LongiBand ="FULL",repet = repet,
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
    adj.r.squared = g$adj.r.squared,
    slope = g$coefficients["abs(ycoord)","Estimate"],
    pval = g$coefficients["abs(ycoord)","Pr(>|t|)"])
    resuAlphaPhylo <- rbind.data.frame(resuAlphaPhylo,interA lphaPhylo)
    for(j in levels(test$LongiBand)){
    petMat <- test[test$LongiBand == j,]
    mod <- lm(mpd~abs(ycoord),data = petMat)
    g <- summary(mod)
    interAlphaPhylo <- cbind.data.frame(age = Age,LongiBand = j,repet = repet,
    adj.r.squared = g$adj.r.squared,
    slope = g$coefficients["abs(ycoord)","Estimate"],
    pval = g$coefficients["abs(ycoord)","Pr(>|t|)"])
    resuAlphaPhylo <- rbind.data.frame(resuAlphaPhylo,interA lphaPhylo)
    }
    }
    write.table(resuAlphaTaxo,"min10OTUs/
AlphaTaxo_LM.txt",sep="\t")
    write.table(resuAlphaPhylo,"min10OTUs/
AlphaPhylo_LM.txt",sep="\t")
    resuAlphaTaxo <- read.table("min10OTUs/
AlphaTaxo_LM.txt",sep="\t")
    resuAlphaPhylo <- read.table("min10OTUs/
AlphaPhylo_LM.txt",sep="\t")
    PerSigni <- function(x){
    sum(x < 0.05)
    }
    # ## Slope with latitude
    resuAlphaTaxo$age = as.numeric(resuAlphaTaxo$age)
    resuAlphaTaxo$age = as.factor(resuAlphaTaxo$age)
    resuAlphaTaxo$signi = resuAlphaTaxo$pval < 0.05
    prop <- table(resuAlphaTaxo$signi,resuAlphaTaxo$age,resu AlphaTaxo$LongiBand)
    prop <- prop[2,,]
    prop <- melt(prop)
    colnames(prop)=c("age","LongiBand","signiProp")
    colo <- blue2red(100)
    colo <- colo[100:1]
    ColorBoxplot <- colo[prop$signiProp]
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
fin <- merge(resuAlphaTaxo,prop)

fin2 <- fin[fin$LongiBand==“FULL”,]

# # Figure not in paper

png(“min10OTUs/AlphaTaxonomic_SlopeWithLatitude_ ThroughTime_min10OTUs.png”,res = 300,width = 1600,heig ht = 1200)

p <- ggplot(fin2,aes(x = age,y = slope,group = age,fill = sign iProp)) + geom_boxplot() + scale_fill_gradient(“significativity (%)”,low = colo[1],high = colo[100]) + theme_ classic() + xlab(“Time [myrs]”) + ylab(“Slope”)

p

dev.off()

# # Figure not in paper

png(“min10OTUs/AlphaTaxonomic_R2WithLatitude_ ThroughTime_min10OTUs.png”,res = 300,width = 1600,heig ht = 1200)

p <- ggplot(fin2,aes(x = age,y = adj.r.squared,group = age,fill = signiProp)) + geom_boxplot() + scale_fill_ gradient(“significativity (%)”,low = colo[1],high = colo[100] ) + theme_classic() + xlab(“Time [myrs]”) + ylab(“Adjusted R^2”)

p

dev.off()

for(i in 1:9){

print(paste(“Time:”,levels(fin2$age)[i],“mean:”,round(mean (fin2$adj.r.squared[fin2$age==levels(fin2$age)[i]]),5), “sd: ”, round(sd(fin2$adj.r.squared[fin2$age==levels(fin$age)[i]]),5), “median: ”, round(median(fin2$adj.r.squared[fin2$age==level s(fin$age)[i]]),5)))

}

# [1] “Time: 0 mean: 0.07574 sd: 0 median: 0.07574”

# [1] “Time: 25 mean: 0.0453 sd: 0.01489 median: 0.04602”

# [1] “Time: 50 mean: 0.03123 sd: 0.01762 median: 0.03232”

# [1] “Time: 75 mean: 0.01788 sd: 0.01669 median: 0.01461”

# [1] “Time: 100 mean: 0.00947 sd: 0.01522 median: 0.00313”

# [1] “Time: 125 mean: 0.00087 sd: 0.00778 median: -0.00193”

# [1] “Time: 150 mean: -0.00161 sd: 0.00412 median: -0.00275”

# [1] “Time: 175 mean: 0.03131 sd: 0.00888 median: 0.03443”

# [1] “Time: 200 mean: 0.05002 sd: 0.01213 median: 0.05802”

fin <- fin[fin$LongiBand!=“FULL”,]

# # Figure 3

png(“min10OTUs/AlphaTaxonomic_ R2WithLatitude_ThroughTime_LongiBand_Min_n10. png”,res = 300,width = 3600, height = 2400)

p <- ggplot(fin,aes(x = age,y = adj.r.squared,group = age,fill = signiProp)) + geom_boxplot() + scale_fill_ gradient(“significativity (%)”,low = colo[1],high = colo[100] ) + theme_classic() + xlab(“Time [myrs]”) + ylab(“Adjusted R^2”) + facet_wrap(~LongiBand)

print(p)
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
dev.off()
# # Figure 3
png("min10OTUs/AlphaTaxonomic_ SlopeWithLatitude_ThroughTime_LongiBand_Min_n10.png",res = 300,width = 3600, height = 2400)
p <- ggplot(fin,aes(x = age,y = slope,group = ag e,fill = signiProp)) + geom_boxplot() + scale_fill_gradient("significativity (%)",low = colo[1],high = colo[100]) + theme_classic() + xlab("Time [myrs]") + ylab("Slope") + fa cet_wrap(~LongiBand)
print(p)
dev.off()
# # AlphaPhylo
resuAlphaPhylo$age = as.numeric(resuAlphaPhylo$age)
resuAlphaPhylo$age = as.factor(resuAlphaPhylo$age)
resuAlphaPhylo$signi = resuAlphaPhylo$pval < 0.05
prop <- table(resuAlphaPhylo$signi,resuAlphaPhylo$age,re suAlphaPhylo$LongiBand)
prop <- prop[2,,]
prop <- melt(prop)
colnames(prop)=c("age","LongiBand","signiProp")
library(colorRamps)
colo <- blue2red(100)
colo <- colo[100:1]
ColorBoxplot <- colo[prop$signiProp]
fin <- merge(resuAlphaPhylo,prop)
fin2 <- fin[fin$LongiBand=="FULL",]
# # Figure not in paper
png("min10OTUs/MPD_SlopeWithLatitude_ ThroughTime_min10OTUs.png",res = 300,width = 1600,heig ht = 1200)
p <- ggplot(fin2,aes(x = age,y = slope,group = age,fill = sign iProp)) + geom_boxplot() + scale_fill_gradient("significativity (%)",low = colo[1],high = colo[100]) + theme_ classic() + xlab("Time [myrs]") + ylab("Slope")
p
dev.off()
# # Figure not in paper
png("min10OTUs/MPD_R2WithLatitude_ThroughTime_ min10OTUs.png",res = 300,width = 1600,height = 1200)
p <- ggplot(fin2,aes(x = age,y = adj.r.squared,group = age,fill = signiProp)) + geom_boxplot() + scale_fill_ gradient("significativity (%)",low = colo[1],high = colo[100] ) + theme_classic() + xlab("Time [myrs]") + ylab("Adjusted R^2")
p
dev.off()
for(i in 1:9){
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
print(paste("Time:",levels(fin2$age)[i],"mean:",round(mean (fin2$adj.r.squared[fin2$age==levels(fin2$age)[i]]),5), "sd: ", round(sd(fin2$adj.r.squared[fin2$age==levels(fin$age)[i]]),5), "median: ", round(median(fin2$adj.r.squared[fin2$age==level s(fin$age)[i]]),5)))

}

#[1] "Time: 0 mean: 0.05967 sd: 0.00121 median: 0.05952" # [1] "Time: 25 mean: 0.05528 sd: 0.03123 median: 0.05031" # [1] "Time: 50 mean: 0.02616 sd: 0.02326 median: 0.02144" # [1] "Time: 75 mean: 0.02144 sd: 0.02529 median: 0.00992" # [1] "Time: 100 mean: 0.04222 sd: 0.04767 median: 0.02195"

# [1] "Time: 125 mean: 0.05844 sd: 0.0405 median: 0.04701"

#[1] "Time: 150 mean: 0.12699 sd: 0.04987 median: 0.12251"

# [1] "Time: 175 mean: 0.05381 sd: 0.0293 median: 0.0498"

#  [1] "Time: 200 mean: 0.03354 sd: 0.0208 median: 0.03201" fin <- fin[fin$LongiBand!="FULL",]

# # Figure 3

png("min10OTUs/MPD_R2WithLatitude_ThroughTime_ LongiBand_min10OTUs.png",res = 300,width = 3600, height = 2400)

p <- ggplot(fin,aes(x = age,y = adj.r.squared,group = age,fill = signiProp)) + geom_boxplot() + scale_fill_ gradient("significativity (%)",low = colo[1],high = colo[100] ) + theme_classic() + xlab("Time [myrs]") + ylab("Adjusted R^2") + facet_wrap(~LongiBand)

print(p)

dev.off()

# # Figure 3

png("min10OTUs/MPD_SlopeWithLatitude_ ThroughTime_LongiBand_min10OTUs. png",res = 300,width = 3600, height = 2400)

p <- ggplot(fin,aes(x = age,y = slope,group = ag e,fill = signiProp)) + geom_boxplot() + scale_fill_ gradient("significativity (%)",low = colo[1],high = colo[100]) + theme_classic() + xlab("Time [myrs]") + ylab("Slope") + fa cet_wrap(~LongiBand)

print(p)

dev.off()
```

(3) The code for generating Figure S1 in this study.

```
##################################

# # Analyses at the genus level

for(z in 1){

################################################ ######

# ## Data loading and preparation for the computation #

################################################ ######

print(paste0("start of replicate ",z))

treeLiver = read.tree("data/changed_phylo_liverwort_la- grange.tree") # Phylogenetic tree at genus level

dataLiver = read.csv("data/Liverworts-database-to-analyse. csv",row.names = 1) # Species matrix (columns = OGUs and rows = species)
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
dataLiver <- dataLiver[,which(apply(dataLiver,2,sum)>=2)] # Restrained the matrix to minimum 2 sp per OGUs
# Remove unused OGUs due to their old names (e.g. Yugoslavia) or the lack of environmental data
aRetirer <- c("X13YUG","X40WHM","X14RUS","X40ASS","X85MST","X85PAT","X85TDF","X90MPE","X21SEL",
"X28STH","X11CZL","X30EVE","X31MGW","X33NCS"," X60GIL","X81SWC","X84BTR")
for(i in 1: length(aRetirer)){
dataLiver <- dataLiver[,colnames(dataLiver)!=aRetirer[i]]
}
# Remove species that is not present in any OGUs
dataLiver <- dataLiver[which(apply(dataLiver,1,sum)!=0),]
# # Matching teh phylogeny with our datasets + Transforming our species data at the genus level
phyloLiver <- read.tree("data/changed_phylo_liverwort_la- grange.tree")
namePhylo <- phyloLiver$tip.label
genera <- unique(sub("_.*","",namePhylo))
ii <- sapply(genera,function(x,y) grep(x,y)
[1],y = namePhylo)
ToRem <- setdiff(1:length(phyloLiver$tip.label),
as.numeric(ii))
phyloLiver<-drop.tip(phyloLiver,c(phyloLiver$tip.
label[ToRem],"Odontoschisma_denudatum_IBC53_USA_ Cephaloziaceae"))
namePhylo <- phyloLiver$tip.label
genera <- sub("_.*","",namePhylo)
phyloLiver$tip.label = genera
speciesLiver = row.names(dataLiver) #Species name
genera = sub("_.*","",speciesLiver)
generaUn <- unique(genera)
dataLiver2 <- as.data.frame(matrix(0,nc = ncol(dataLiver), nr = length(generaUn)))
colnames(dataLiver2) = colnames(dataLiver)
rownames(dataLiver2) = generaUn
for(j in 1:length(generaUn)){
test <- apply(dataLiver[genera == generaUn[j],],2,function(x)
{as.numeric(sum(x) > 0)})
dataLiver2[generaUn[j],] = test
}
dataLiver <- dataLiver2
speciesLiver = row.names(dataLiver) #Species name commLiver<-t(dataLiver) #transposition of the dataset FilterData <- match.phylo.comm(phyloLiver,commLiver) phyloLiver <- FilterData$phy
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
commLiver <- FilterData$comm
# Splitting the OGUs into different regions #
tabGroupeLiver <- read.csv("data/tableau-centro-trop2. csv") # Contains OGU center coordinates with the delimitation T/ETN/ETS
aSupp <-c() # vector that will contain position of OGUs that is not used in the centroid dataset
# Give position of unused OGUs
for(i in 1: length(as.character(tabGroupeLiver[,3]))){
if(sum(colnames(dataLiver) == paste0("X",as.character(tab
GroupeLiver[i,3]))) == 0){
aSupp <- c(aSupp,i)
}
}
# Remove it from the table
tabGroupeLiver<-tabGroupeLiver[-aSupp,]
rownames(tabGroupeLiver) = tabGroupeLiver$a.garder
# # Adaptation for computation through time
phyloLiver$node.label = c(1:phyloLiver$Nnode) + length(p hyloLiver$tip.label) #add labels to node matching their number in object phylo
nodeagesLiver <- branching.times(phyloLiver) #min age of branch starting up to node
# add abundances up the phylogeny #
commwnodesLiver<-as.data.frame(commLiver)
commwnodesLiver[,length(commLi
ver[1,]) + (1:phyloLiver$Nnode)] = 0
colnames(commwnodesLiver[,(length(co mmLiver[1,]) + (1:phyloLiver$Nnode))])<- length(commLiver[1,]) + c(1:phyloLiver$Nnode)
# identify node paths from root to each tip #
npLiver<-nodepath(phyloLiver)
# compute abundance for each node #
for(i in 1:length(commLiver[,1])){
for(t in 1:length(phyloLiver$tip.
label)) if(commwnodesLiver[i,t] > 0)
commwnodesLiver[i,npLiver[[t]][1:(length(npLiver[[t]])-1)]]
= commwnodesLiver[i,npLiver[[t]][1:(length(npLiver[[t]])-1)]
] + commwnodesLiver[i,npLiver[[t]][length(npLiver[[t]])]]
}
write.table(commwnodesLiver,"V2/commwnodes. txt",sep="\t")
# compute inf/sup age of each branch #
agenLiver = as.data.frame(matrix(ncol = 2,nrow = length(co
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
mmwnodesLiver[1,])))
    colnames(agenLiver)=c("sup","inf")
    edgeLiver<-rbind(phyloLiver$edge,c((length(phyloLiver$ tip.label) + 1),(length(phyloLiver$tip.label) + 1))) #make an edge table including the root
    edgeLiver<-edgeLiver[order(edgeLiver[,2]),] #ordered phylo$edge by descendent tip/node number
    agenLiver$inf[1:length(commLiver[1,])] = 0 #inf age of tips = 0
    agenLiver$sup[] = nodeagesLiver[edgeLiver[,1]- length(phyloLiver$tip.label)]
    agenLiver$inf[(length(phyloLiver$tip.label) + 1):length(agenLiver$inf)] = nodeagesLiver
    write.table(agenLiver,"v2/age_nodes_inf_sup.txt",sep="\t") #write this branch age table
    # Dataset preparation for the computation through time
    dataPlot <- as.data.frame(matrix(0,ncol = length(phyloLiver$tip.label) + 3,nr = nrow(agenLiver)))
    dataPlot[1] = 1:nrow(agenLiver) # name of the node
    colnames(dataPlot)[1]="NodeAge"
    colnames(dataPlot)[4:(length(phyloLiver$tip.label) + 3)] = phyloLiver$tip.label
    colnames(dataPlot)[2:3]=c("ageSup","ageInf")
    dataPlot[,2:3] = agenLiver #Age of the node
    # species present for each node #
    for(i in 1:nrow(dataPlot)){
    spNode <- phyloLiver$tip.label[sapply(1:length(phyloLiver$tip.label),present,data = npLiver,i = i)==1]
    dataPlot[i,spNode] = 1
    }
    maxi <- max(unlist(lapply(npLiver,length)))
    groupeChemin <- sapply(1:maxi,noeudListe,liste = npLiver) #dataframe containing all the species path in the tree
    nomSp <- colnames(dataPlot[,-c(1:3)])
    PhyloTime <- phyloLiver
    dataTime <- commLiver
#     # Create a matrix where node is present per OGU numberNode <- nrow(phyloLiver$edge)
for(i in 1:(numberNode + 1)){
    noeud = i
    noeudPresent<-lapply(lapply(npLiver,function(x,noeud) {if(length(which(x ==
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
noeud))!=0){return(x)}},noeud = noeu d),length)
    nomSpRegroup <- phyloLiver$tip.
label[which(noeudPresent > 0)]
    petitTab <- commLiver[,nomSpRegroup]
    if(is.matrix(petitTab)){
    colARajouter<- apply(petitTab,1,sum)
    colARajouter[colARajouter > 1] = 1
    }
    else{colARajouter <- petitTab}
    dataTime <- data.frame(dataTime,colARajouter)
    colnames(dataTime)[colnames(dataTime)==“colARajou ter”] = noeud
    }
    dataTime <- dataTime[,-c(1:length(colnames(commLiver)))]
#   Remove species only keep node name
    colnames(dataTime)=sub(pattern = “X”,replacement = “”,colnames(dataTime))
    write.table(dataTime,“v2/commLiveramongNodes. txt”,sep=“\t”)
    tabledivLiverTime = specnumber(dataTime)
    rm(nodeagesLiver,colARajouter,noeudPresent,petitTab,grou
peChemin,numberNode,edgeLiver,commwnodesLiver,npLiver ,maxi,spNode)
    ################################################
######
    ####################
    # # Temporal loop ##
    ###################
    resuAlphaTaxo <- NULL
    resuAlphaPhylo <- NULL
    PhyloTime<-phyloLiver
    nLineage <- c()
    for(i in 1:201){
    age = i-1 #began at time = 0 to 200 Mya
    if(i == 1){age = 0.0000000001}
    # Node Selection based on time
    commLiverTime<-dataPlot[dataPlot$ageInf < age & dataPlot$ageSup>=age,-c(1:3)]
    # regroup all the species that are in the same node to consider it as only one “species”
    nombreSpAregrouper <- which(apply(commLiverTime,1, sum)>=2)
    dataSP <- apply(commLiverTime,1,estEgaleA1)
    spName <- do.call(rbind,lapply(dataSP,nomASuppr,nom Col = nomSp))
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
spNameToChange <- spName[,1]
dataTime2 <-dataTime[,rownames(commLiverTime)] #
Only keeps species that will be used for the computation
colnames(dataTime2) = spNameToChange
if(length(nombreSpAregrouper)!=0){#Regroup some species When we go through time
for(j in 1: length(rownames(commLiverTime))){
node <- which(phyloLiver$edge[,1]==as.
numeric(rownames(commLiverTime))[j])
if(length(node)!=0){
PhyloTime$edge.length[node] = 0 #decrease the previous branch to 0
}
}
}
b <- table(spNameToChange)
ToMerge <- names(b)[b > 1]
if(length(ToMerge) > 0){
for(g in 1:length(ToMerge)){
SeveralCol <- apply(dataTime2[,grep(ToMerge[g],colnames (dataTime2))],1,sum)
dataTime2 <- dataTime2[,-grep(ToMerge[g],colnames(data
Time2))]
dataTime2 <- cbind.data.frame(dataTime2,SeveralCol)
colnames(dataTime2)[ncol(dataTime2)] = ToMerge[g]
}
}
dataTime2<- dataTime2[which(apply(dataTime2,1,sum) > 1),]  # Keeps
OGUs that have at least 1 "species"
dataTime2<-dataTime2[,which(apply(dataTime2,2,
sum) > 0)] # Keeps OGUs that have at least 1 "species"
if(i==1){taillechgt <-c(i,nrow(dataTime2))}
if(i!=1){taillechgt <-rbind(taillechgt,c(i,nrow(dataTime2)))}
if(i %in% seq(1,201,by = 25)){
print(age)
# Remove unused species
PhyloTime2 <- keep.tip(PhyloTime,colnames(dataTime2))
tabledivLiverTime2 <- specnumber(dataTime2)
nLineage = c(nLineage,length(PhyloTime2$tip.label))
write.table(tabledivLiverTime2,paste0("V2/
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
AlphaTaxo_",z,"_time",(i-1),".txt"),sep="\t")
    lati <- tabGroupeLiver[sub("X","",names(tabledivLiverT ime2)),]
    lati <- lati$ycoord
    mod <- lm(tabledivLiverTime2~abs(lati))
    g <- summary(mod)
    interAlphaTaxo <- cbind.data.frame(age=(i-1),adj.r.squared = g$adj.r.squared,slope = g$coefficients["abs(lati)","Estim ate"], pval = g$coefficients["abs(lati)","Pr(>|t|)"])
    resuAlphaTaxo <- rbind.data.frame(resuAlphaTaxo,interAlp haTaxo)
    dis <- cophenetic(PhyloTime2)
    alphaPhylo <- mpd(dataTime2,dis)
    names(alphaPhylo) = names(tabledivLiverTime2)
    write.table(alphaPhylo,paste0("V2/AlphaPhylo_",z,"_
time",(i-1),".txt"),sep="\t")
    mod <- lm(alphaPhylo~abs(lati))
    g <- summary(mod)
    interAlphaPhylo <- cbind.data.frame(age=(i-1),adj.r.squared = g$adj.r.squared,slope = g$coefficients["abs(lati)","Estim ate"], pval = g$coefficients["abs(lati)","Pr(>|t|)"])
    resuAlphaPhylo <- rbind.data.frame(resuAlphaPhylo,interA lphaPhylo)
    }
    }
    write.table(resuAlphaPhylo,paste0("V2/LatiCorr_ AlphaPhylo_",z,".txt"),sep="\t")
    write.table(resuAlphaTaxo,paste0("V2/LatiCorr_
AlphaTaxo",z,".txt"),sep="\t")
    write.table(nLineage,paste0("V2/NLineages.txt"),sep="\t")
    print(paste0("end of the replicat ",z))
    }
    ##################
    # # MPD~Latitude for 3 longitudinal band
    ##############################
    tabGroupeLiver <- read.csv("data/tableau-centro-trop2. csv") # Contains OGU center coordinates with the delimitation T/ETN/ETS
    tabGroupeLiver$LongiBand = "0"
    tabGroupeLiver$LongiBand[tabGroupeLiver$xco ord < (-60)] = "Americas"
    tabGroupeLiver$LongiBand[tabGroupeLiver$xcoord > 60] = "Asia/Oceania"
    tabGroupeLiver$LongiBand[tabGroupeLiver$xcoord <= 60
& tabGroupeLiver$xcoord >= (-60)] = "Europe/Africa"
    tabGroupeLiver$a.garder = paste0("X",tabGroupeLiver$a. garder)
    filesAlpha<-list.files(pattern="AlphaTaxo_",path="V2",full. names = T)
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
resuAlphaTaxo <- NULL
for(i in 1:length(filesAlpha)){
Age <- sub(".*time","",filesAlpha[i])
Age <- sub(".txt","",Age,fixed = T)
repet <- sub("_time.*","",filesAlpha[i])
repet <- sub("outputs/AlphaTaxo_","",repet,fixed = T)
inter <- cbind.data.frame(read.
table(filesAlpha[i],sep="\t"),Age,repet)
colnames(inter)[1]="mpd"
inter$OGU = row.names(inter)
row.names(inter) = NULL
tab2 <- tabGroupeLiver[tabGroupeLiver$a.garder %in% inter$OGU,]
test <- merge(inter,tab2,by.x="OGU",by.y= "a.garder")
test$LongiBand = as.factor(test$LongiBand)
for(j in levels(test$LongiBand)){
petMat <- test[test$LongiBand == j,]
mod <- lm(mpd~abs(ycoord),data = petMat)
g <- summary(mod)
interAlphaTaxo <- cbind.data.frame(age=Age,LongiBand=j ,repet=repet,adj.r.squared =
g$adj.r.squared,slope = g$coeffici ents["abs(ycoord)","Estimate"], pval = g$coefficients["abs(yc
oord)","Pr(>|t|)"])
resuAlphaTaxo <- rbind.data.frame(resuAlphaTaxo,interAlp haTaxo)
}
}
resuAlphaTaxo$age = as.numeric(resuAlphaTaxo$age)
resuAlphaTaxo$age = as.factor(resuAlphaTaxo$age)
resuAlphaTaxo$signi = resuAlphaTaxo$pval < 0.05
plyr::ddply(resuAlphaTaxo,.(age,LongiBand),summarise,M ean = mean(adj.r.squared))
prop <- table(resuAlphaTaxo$signi,resuAlphaTaxo$age,resu AlphaTaxo$LongiBand)
colnames(prop)=c("age","LongiBand","signiProp")
fin <- resuAlphaTaxo
# # Fig S1
colo <- c("#863333","#57A29A")
png("V2/AlphaTaxonomic_R2WithLatitude_ ThroughTime_LongiBand.png",res = 300,width =
3600, height = 2400)
p <- ggplot(fin,aes(x = age,y = slope,group = age,col our = signi)) + geom_point(aes(size =
adj.r.squared)) + scale_colour_discrete("Significant",type = colo) + the
me_classic() + xlab("Time [myrs]") + ylab("Slope") + fac et_wrap(~LongiBand)
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
print(p)
dev.off()
filesAlpha <- list.files(pattern = "AlphaPhylo_",path = "V2",full.names = T)
filesAlpha <- filesAlpha[-grep("LatiCorr",filesAlpha)]
resuAlphaPhylo <- NULL
for(i in 1:length(filesAlpha)){
Age <- sub(".*time","",filesAlpha[i])
Age <- sub(".txt","",Age,fixed = T)
repet <- sub("_time.*","",filesAlpha[i])
repet <- sub("outputs/AlphaPhylo_","",repet,fixed = T)
inter <- cbind.data.frame(read.
table(filesAlpha[i],sep="\t"),Age,repet)
colnames(inter)[1]="mpd"
inter$OGU = row.names(inter)
row.names(inter) = NULL
tab2 <- tabGroupeLiver[tabGroupeLiver$a.garder %in% inter$OGU,]
test <- merge(inter,tab2,by.x="OGU",by.y= "a.garder")
test$LongiBand = as.factor(test$LongiBand)
for(j in levels(test$LongiBand)){
petMat <- test[test$LongiBand == j,]
mod <- lm(mpd~abs(ycoord),data = petMat)
g <- summary(mod)
interAlphaPhylo <- cbind.data.frame(age=Age,LongiBand= j,repet=repet,adj.r.squared =
g$adj.r.squared,slope = g$coeffici ents["abs(ycoord)","Estimate"], pval = g$coefficients["abs(yc
oord)","Pr(>|t|)"])
resuAlphaPhylo <- rbind.data.frame(resuAlphaPhylo,interA lphaPhylo)
}
}
resuAlphaPhylo$age = as.numeric(resuAlphaPhylo$age)
resuAlphaPhylo$age = as.factor(resuAlphaPhylo$age)
resuAlphaPhylo$signi = resuAlphaPhylo$pval < 0.05
fin <- resuAlphaPhylo
# # Fig S1
png("V2/MPD_R2WithLatitude_ThroughTime_ LongiBand.png",res = 300,width = 3600, height =
2400)
p <- ggplot(fin,aes(x = age,y = slope,group = age,col our = signi)) + geom_point(aes(size =
adj.r.squared)) + scale_colour_discrete("Significant",type = colo) + the me_classic() + xlab("Time
[myrs]") + ylab("Slope") + fac et_wrap(~LongiBand)
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
print(p)
dev.off()
```

(4) The code for generating Table 1 in this study.

```
dat = read.csv("SLM_withoutNA175_200.csv",header = T) ##change file by different intervals (without NA value), 9 intervals in total
dat = dat[,-c(1)]
source("R script_5_SAR_Functions.r")
# ---------- NEIGHBOURHOOD STRUCTURE ---------- #
#Neighbourhood
coords <- as.matrix(cbind(dat$longCentre,dat$latCentre))
library(spdep)
nb <- knearneigh(coords,k = 3, longlat = TRUE) # k = 1
nb1 <- knn2nb(nb)
# nb1 <- dnearneigh(coords,0,2000,longlat = TRUE)
plot(nb1, coords)
# # Moran's I
nb2 <- nb2listw(nb1, glist = NULL, style="W", zero. policy = FALSE)
# ---------- Species Richness ----------#
colnames(dat)
env <- dat[,c(5,6,9,10)]
sar_SR <- AllSubsetModelSelection(X = env, Y = dat$MPD200, nbw = nb2) #change $MPD for different intervals
##################################################
##########
##################################################
##########
##################################################
##########
# Returns: A list with five elements:
# Model.AIC.scores: vector of AIC scores for each subset model
# Model.AIC.weights: vector of AIC weights for each model
# Variables: a data frame with one row for each model, one column for each
# predictor. Contains a 1 if that predictor was used in that model
# and a zero otherwise
# Best.Model: the variables included in the model with the lowest AIC
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
    # Coefficients: a data frame containing the OLS fitted coefficients for
    # each variable included in each model. Contains an NA when the
    # variable was not included in the model (corresponds with 0 values
    # in the Variables matrix).
    sar_SR$Best.Model
    ################################################ ############
    # Calculate summed AIC weights for predictor variables, using the
    # output of AllSubsetModelSelection
    # Arguments:
    # MS: A list object containing the output of the AllSubsetsModelSelection() function
    # Returns: A vector with summed AIC weights for each predictor variable
    sar_sumW <- SumWeights(sar_SR)
    sar_sumW <- data.frame(variable = colnames(env), sumW = sar_sumW, id = 1:length(sar_sumW))
    ################################################
###########
    # # Run again for the best model
    sar_SR$Best.Model
    model = lm(MPD200~bio12 + tempVelo + precVelo,data = dat)#change individual variables by best model
    library(spatialreg)
    sar_SR.best <- errorsarlm(model,listw = nb2,na.action = na. omit,data = dat)
    h <- dat$MPD200 # the response variable
SLM = c(sar_SR.best$coefficients[2]*sd(dat$bio12)/sd(h),
sar_SR.best$coefficients[3]*sd(dat$tempVelo)/sd(h),
sar_SR.best$coefficients[4]*sd(dat$precVelo)/sd(h),
    R.squared = cor(x = fitted(sar_SR.best), y = h,
method="pearson")^2,
    AIC = AIC(sar_SR.best),
    moran. I = moran.test(residuals(sar_SR.best), listw = nb2)$estimate[1],
    moran.I.p.value = moran.test(residuals(sar_SR.best), listw = nb2)$p.value)
    names(SLM)[names(SLM)=="moran.I.Moran I
statistic"]="moran.I"
    SLM <- data.frame(variable = names(SLM), val = as. numeric(SLM),pValue = NA)
    SLM$pValue[1:(length(sar_SR.best$coefficients)-1)]
<- as.numeric(summary(sar_SR.best)$Coef[-1,4])
    # # merge
    SLM <- merge(SLM, sar_sumW, by="variable", all = TRUE)
    SLM <- SLM[order(SLM$id),]
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
    SLM$pred="SR"
    SLM$val <- round(SLM$val, 3)
SLM$pValue <- round(SLM$pValue,4) # save
    write.csv(SLM, "AllMPD200.csv", row.names = F)
    The following content is the code for "R script_5_SAR_Functions.r".
    # # The original R script is from Brody Sandel.
    ################################################
#########
    # All subsets model selection using AIC
    # Take a response variable (Y) and a matrix of predictor
    # variables (X). Using OLS, take all subsets of possible
    # predictor variables from X, run the regression and save
    # AIC scores.
    #
    # This can be used for model selection, and also for model
    # averaging (multi-model inference)
    # Arguments:
    # X: a matrix or data frame of predictor variables
    # Y: a vector containing the response variable
    # Returns: A list with five elements:
    # Model.AIC.scores: vector of AIC scores for each subset model
    # Model.AIC.weights: vector of AIC weights for each model
    # Variables: a data frame with one row for each model, one column for each
    # predictor. Contains a 1 if that predictor was used in that model
    # and a zero otherwise
    # Best.Model: the variables included in the model with the lowest AIC
    # Coefficients: a data frame containing the OLS fitted coefficients for
    # each variable included in each model. Contains an NA when the
    # variable was not included in the model (corresponds with 0 values
    # in the Variables matrix).
    # #####model for selection
    AllSubsetModelSelection = function(X,Y,nbw)
    {
    require(gtools)
    ####Feng's SAR
    require(spdep)
    require(spatialreg)
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
require(ncf)
nv = ncol(X)
d = data.frame(Y,X)
#define all possible models with a big matrix
modelmatrix = matrix(0,nrow = 0,ncol = nv)
for(i in 1:nv)
{
comb = combinations(nv,i)
mat = matrix(0,nrow = nrow(comb),ncol = nv)
for(j in 1:nrow(comb))
{
mat[j,comb[j,]] = 1
}
modelmatrix = rbind(modelmatrix,mat)
}
print(paste("Fitting",nrow(modelmatrix),"models. This
could take a while"))
# write.csv(modelmatrix, "modelmatrix.csv")
#evaluate all possible models
aic = rep(NA,nrow(modelmatrix))
coefs = matrix(NA,nrow = nrow(modelmatrix),ncol = ncol( modelmatrix))
for(i in 1:nrow(modelmatrix)){
varnames = colnames(X)[which(modelmatrix[i,] == 1)]
form = paste("Y~",paste(varnames,collapse="+"),sep = "")
model = lm(as.formula(form),data = d)
# ### Feng's SAR
sar.model = errorsarlm(model,listw = nbw,na.action = na.
omit,data = d)
coefs[i,which(modelmatrix[i,]      ==              1)] = sar.
model$coefficients[2:length(sar.model$coefficients)]
aic[i] = AIC(sar.model)
}
#calculate best AIC score and AIC weights for each model
best = which(aic == min(aic))
bestaic = aic[best]
temp = aic-bestaic
temp2 = exp(-1*temp/2)
```

Published in : Annals of Botany (2025), mcaf051
DOI: 10.1093/aob/mcaf051
Status : Postprint (Author's version)

```
    aicweights = temp2/sum(temp2)
    coefsframe = data.frame(coefs)
    modelframe = data.frame(modelmatrix)
    colnames(modelframe) = names(X)
    colnames(coefsframe) = names(X)
    output = list(aic,aicweights,modelframe,modelframe[best,], coefsframe)
    names(output) = c("Model.AIC.scores","Model.AIC.
weights","Variables","Best.Model","Coefficients")
    return(output)
    }
    ################################################
###########
    # Calculate summed AIC weights for predictor variables, using the
    #   output of AllSubsetModelSelection
    #   Arguments:
    # MS: A list object containing the output of the AllSubsetsModelSelection() function
    # Returns: A vector with summed AIC weights for each predictor variable
    SumWeights = function(MS)
{
summedWeight = rep(NA,ncol(MS$Variables))
for(i in 1:ncol(MS$Variables))
{
index = which(MS$Variables[,i] == 1)
summedWeight[i] = sum(MS$Model.AIC.weights[index]) }
return(summedWeight)
```