# Updated angiosperm family tree for analyzing phylogenetic diversity and community structure

Markus Gastauer[1,2]* and João Augusto Alves Meira-Neto[2]

## ABSTRACT

The computation of phylogenetic diversity and phylogenetic community structure demands an accurately calibrated, high-resolution phylogeny, which reflects current knowledge regarding diversification within the group of interest. Herein we present the angiosperm phylogeny R20160415.new, which is based on the topology proposed by the Angiosperm Phylogeny Group IV, a recently released compilation of angiosperm diversification. R20160415.new is calibratable by different sets of recently published estimates of mean node ages. Its application for the computation of phylogenetic diversity and/or phylogenetic community structure is straightforward and ensures the inclusion of up-to-date information in user specific applications, as long as users are familiar with the pitfalls of such hand-made supertrees.

**Keywords:** angiosperm diversification, APG IV, community tree calibration, megatrees, phylogenetic topology

## Introduction

The phylogenetic structure of a biological community determines whether species that coexist within a given community are more closely related than expected by chance, and is essential information for investigating community assembly rules (Kembel & Hubbell 2006; Gastauer & Meira-Neto 2014a; Miazaki *et al.* 2015; Lamare *et al.* 2016) as well as determining the evolutionary processes that generated extant biodiversity (Fine & Kembel 2011; Gastauer *et al.* 2015a). More recently, the use of phylogenetic diversity to describe the amount of evolutionary history represented within a sample has gained importance as an indicator for conservation purposes (Forest *et al.* 2007; Huang *et al.* 2016; Arponen & Zupan 2016). The correct computation of these measures demands an accurately calibrated high-resolution phylogeny comprising the entire taxonomic group under study (Gastauer & Meira-Neto 2013).

The constant increase in knowledge about the phylogenetic relationships among taxa (e.g., Cox *et al.* 2014) requires regular revision of applied phylogenies in order to incorporate novel data and avoid out-dated information in analyses of phylogenetic diversity and community structure. For vascular plants, calibratable phylogenies (i.e., Gastauer & Meira-Neto 2016) are based on APG III (2009), nevertheless recent advances in angiosperm phylogeny (i.e., APG IV 2016) have made it necessary to update them.

Therefore, the aim of this study is to provide a fully resolved, up-to-date angiosperm family tree based on APG IV (2016) and Stevens (2016) including features necessary for its accurate calibration. Such a tree will permit the inclusion of recent advances regarding angiosperm phylogeny in user-specific analyses of phylogenetic diversity and phylogenetic community structure.

[1] Instituto Tecnológico Vale, Rua Boaventura da Silva, 955, 66055-090, Belém, PA, Brazil
[2] Laboratório de Ecologia e Evolução de Plantas, Universidade Federal de Viçosa, 36570-000, Viçosa, MG, Brazil

* Corresponding author: markus.gastauer@itv.org

## Materials and methods

### Tree topology

For our angiosperm family tree we used the Newick format, which is required by most tools used for computing phylogenetic community structure or calculating phylogenetic diversity. In contrast to the NEXUS format, the Newick format is fully compatible with Phylocom 4.2 (Webb *et al.* 2002); Newick files can be imported straightforward within the R environment (R Core 2016) using the 'read.tree' command from the 'picante' package (Kembel *et al.* 2010).

The backbone of our fully resolved angiosperm phylogeny is based on APG IV (2016). Phylogenetic relationships among all the angiosperm orders recognized by this updated classification scheme were imported from this publication. Family relationships within orders were acquired from Stevens (2016) and inserted into the backbone with two exceptions. First, we borrowed the order phylogeny for Cucurbitales from Filipowiz & Renner (2010), because they place Apodanthaceae within Cucurbitales as suggested by APG IV (2016), while this family is missing in Stevens' Cucurbitales phylogeny. Second, we adopted Xiang *et al.*'s (2011) phylogeny of Cornales because with its posterior probabilities from Bayesian analysis being larger than 90 %, it offers higher support for interfamilial nodes than Stevens (2016). Nevertheless, we acknowledge that position of Hydrostachyaceae remains doubtful (Magallón *et al.* 2015).

Some family names that are still recognized as legitimate by the Missouri Botanical Garden (2016) are pooled within others in APG IV (for details, see Tab. 1). Nevertheless, automated name and classification checking services such as the Taxonomic Name Resolution Service (TNRS, Boyle *et al.* 2013) still return these out-of date classifications. Therefore, we included them within the family tree at the phylogenetic positions as indicated by Stevens (2016). To indicate their status as families that are no longer accepted by APG IV, we labeled them with the suffix '_NA'. This procedure allows their usage, but compels a manifestation by the user that they are referring to former classifications.

All internal nodes within our family tree were labeled. The node representing the most recent common ancestor of a well-known clade receives its name. This includes all families and orders as well as higher-level classifications such as fabids, rosids, eudicots, monocots and magnoliids. All other nodes were labeled with names that included the extreme positions of all the descendants of the next level, combined by the word "to" (i.e., the clade [[Joinvilleaceae + Ecdeiocoleaceae] + Poaceae] received the name 'joinvilleaceae_to_poaceae')

### 'ages' files for R20150415.new calibration

Two recent comprehensive studies about angiosperm diversification times are available in the literature (i.e., Bell *et al.* 2010; Magallón *et al.* 2015). Mean age estimates for corresponding nodes between the topology of R20160415.new and the phylogenies proposed within these studies were compiled in 'ages' files for the calibration of the megatree using the branch length adjustment (bladj) algorithm from the Phylocom-4.2 package. Since the bladj algorithm calibrates the phylogeny by dating internal nodes with unique values and distributing un-dated nodes evenly between dated nodes, different mean age estimates from exponential (BEAST[a]) or lognormal (BEAST[b]) distributions (Bell *et al.* 2010), as well as those resulting from penalized likelihood (PL) or uncorrelated lognormal (UCLN) methods (Magallón *et al.* 2015), resulted in four different calibration sets available as different 'ages' files: ages_bell_exp.txt, ages_bell_logn.txt, ages_magallon_PL.txt and ages_magallon_UCLN.txt (see Tab. S1 in supplementary material for a complete list of mean node estimations and their standard deviations).

When superior nodes were estimated to be the same mean age or even younger than their descendants, the bladj algorithm is only able to date the older node correctly. The younger node(s) are then distributed equally between the older, dated node and the subsequent node containing age estimates, thus distorting the calibrated trees. In order to avoid this, we altered the age of the descendant node by -0.01 Myr, because this will ensure the maintenance of a topology with less influence on measures of phylogenetic diversity or phylogenetic community structure than would distortions caused by equal node distribution (Gastauer & Meira-Neto 2016). When three subsequent nodes were estimated to be

**Table 1.** Families in R20160415.new that are not included in APG IV, and their phylogenetic positions as indicated by APG III (2009).

| Order | APG IV family | Families from former classification systems pooled to APG IV family |
|---|---|---|
| Piperales | Aristolochiaceae | Asaraceae, Hydnoraceae, Lactoridaceae |
| Poales | Restionaceae | Anarthriaceae, Centrolepidaceae |
| Asparagales | Asphodelaceae | Xanthorrhoeaceae |
| Geraniales | Francoaceae | Vivianiaceae, Greyiaceae, Melianthaceae |
| Santalales | Olacaceae (not monophyletic) | Aptandraceae, Coulaceae, Erythropalaceae, Octoknemaceae, Strombosiaceae, Ximeniaceae |
|  | Santalaceae | Amphorogynaceae, Cervantesiaceae, Comandraceae, Nanodeaceae, Thesiaceae, Viscaceae |

the same mean age, we corrected the age of the superior node by +0.01 Myr. If a superior node was estimated to be younger than its descendants, the age estimate of the superior node was removed from the 'ages' file.

*Proof of concept*

The tree topology and the calibration were applied to two available datasets. The Forest of Seu Nico Forest Dynamics Plot (FSN) dataset from the municipality of Viçosa, Minas Gerais, Brazil, describes trees that occur within a one-hectare plot that is divided into 100 10 m x 10 m subplots (Gastauer & Meira-Neto 2014b; Gastauer *et al.* 2015b; c). A discussion of outcomes of phylogenetic community structure analyses from FSN may be found in Gastauer & Meira-Neto (2014a). The Eifel Grassland dataset comprises the occurrences of species in 62 plots of 1 m$^2$ from different grassland communities from the Eifel in North Rhine-Westphalia, Germany (M Gastauer unpubl. res.).

For proof of concept, we checked the family-level classification of all angiosperm species from both datasets with the TNRS (Boyle *et al.* 2013). Then, we inserted them, according to their family classification, into R20160415. new by the phylomatic function in Phylocom 4.2 (Webb & Donoghue 2005). The resulting community trees were calibrated using the Phylocom's bladj algorithm in combination first with ages_bell_exp.txt or, during the second calibration, with ages_magallon_UCLN.txt; then, the Mean Pairwise Distance (MPD), the Mean Nearest Taxon Distance (MNTD), the Net Relatedness Index (NRI), the Nearest Taxon Index (NTI, Webb *et al.* 2002) and Faith's Phylogenetic Diversity (PD, Faith 1992) were computed for each plot and subplot using Phylocom 4.2. Additionally, the standard effect size of the PD (ses.PD) was computed in the R environment. To compute the NRI, the NTI and the ses.PD, we randomized the species composition of each plot and subplot 999 times using the complete phylogeny pool of each dataset.

## Results and discussion

The resulting angiosperm family tree (S2 in supplementary material) was called R20160415.new due to its high resolution (R) containing only branches with confidence levels (Bootstrap values or posterior probabilities from Bayesian analysis) larger than 80 % and its release date of April 15, 2016. It included all the 64 orders and 416 families recognized by APG IV (2016). As in APG IV, we used Asteraceae (not Compositae), Fabaceae (not Leguminosae), Poaceae (not Gramineae), Apiaceae (not Umbelliferae), Arecaceae (not Palmae), Brassicaceae (not Cruciferae), Clusiaceae (not Guttiferae) and Lamiaceae (not Labiatae). Authors who wish to use the traditional names should change them in the plain text archive of R20160415.new.

To match all the families that are still recognized by the Missouri Botanical Garden but not by APG IV, we maintained 21 family names from former classifications, and added the suffix '_NA' (Tab. 1). Furthermore, because the position of *Peltanthera* was unclear, it was placed as sister to the [Calceolariaceae + Gesneriaceae] clade as proposed by Stevens (2016). Therefore, R20160415.new contains 438 terminals and 402 fully labeled internal nodes.

Differences to its precursor R20120829mod.new (Gastauer & Meira-Neto 2016) are the up-to-date placement of the newly recognized orders Dilleniales, Metteniusales, Icacinales, Boraginales and Vahliales and their subordered families (APG IV 2016). Furthermore, the recently described families Maundiaceae, Apodanthaceae, Peraceae, Ixonanthaceae, Francoaceae, Petenaeaceae, Macarthuriaceae, Microteaceae, Kewaceae, Petiveriaceae and Mazaceae are included, permitting the straightforward incorporation of species from these families within R20160415.new.

Although APG IV comprises more families than APG III, R20160415.new contains fewer terminal nodes than its antecessor R20120829mod.new (Gastauer & Meira-Neto 2016). This is because the latter comprises a complete euphyllophyte phylogeny with 37 monilophyte and 13 gymnosperm families. For researchers interested in these groups, we updated the euphyllophyte phylogeny with APG IV, which is available as R20160415_euphyllophyte.new (S3 in supplementary material). Furthermore, R20120829mod. new contains phylogenetic information about genera and/ or species from 18 families. We withdrew these intrafamilial topologies because they do not comprise the complete phylogeny of the families; they were often restricted to a few of hundreds of genera or species and therefore counterfeit a precision that was not actually provided. Nevertheless, fully functional trees containing this information are available as R20160415_families.new (angiosperms only, S4 in supplementary material) and R20160415_euphyllophyte_ families.new (complete euphyllophyte, S5 in supplementary material).

By comparing the topology of R20160415.new with the dated phylogenies from literature, we identified 267 nodes corresponding to nodes of Magallón's tree (Magallón *et al.* 2015) as well as 306 nodes corresponding to nodes of Bell's phylogeny (Bell *et al.* 2010, Tab. S1 in supplementary material). A few of the mean age estimates were misleading (Tab. 2), therefore, the 'ages' files contain mean age estimates for only 304 nodes in ages_bell_exp. txt (S6 in supplementary material) and 302 nodes in ages_bell_logn.txt (S7 in supplementary material), while ages_magallon_PL.txt (S8 in supplementary material) and ages_magallon_UCLN.txt (S9 in supplementary material) compile mean age estimates for 267 nodes.

Bell *et al.* (2010) provide a greater number of crown age estimates for angiosperm families than Magallón *et al.* (2015), although these might be biased towards erroneously young ages in heterogeneous measures (Magallón *et al.* 2015).

**Table 2.** Corrections to the age estimates of the corresponding nodes in 'ages' files due to misleading information in Bell *et al.* (2010).

| Node names | Age (BEAST[a]) | Age in ages_bell_exp.txt | Age (BEAST[b]) | Age in ages_bell_logn.txt |
|---|---|---|---|---|
| magnoliales_to_asterales | 130 | **130.01** | 146 | **146.01** |
| poales_to_asterales | 130 | 130 | 146 | 146 |
| monocots | 130 | **129.99** | 146 | **145.99** |
| eudicots | 129 | 129 | 129 | 129 |
| proteales_to_asterales | 126 | 126 | 129 | **128.99** |
| trochandrales_to_asterales | 121 | 121 | 134 | **Removed** |
| gunnerales_to_asterales | 117 | **Removed** | 121 | **Removed** |
| superrosids_to_superasterids | 119 | 119 | 123 | **Removed** |
| superrosids | 117 | 117 | 128 | 128 |
| rosids | 101 | **Removed** | 125 | 125 |
| malvids_to_fabids | 109 | 109 | 111 | 111 |
| rosales_to_fagales | 96 | 96 | 96 | 96 |
| cucurbitales_to_fagales | 96 | **95.99** | 96 | **95.99** |
| campanulids | 93 | 93 | 100 | 100 |
| asterales_to_paracryphiales | 93 | **92.99** | 100 | **99.99** |
| goodeniaceae_to_asteraceae | 44 | 44 | 40 | **Removed** |
| asteraceae | 40 | 40 | 43 | 43 |

Comparing 'ages' files from both publications, we found age estimates from 154 nodes to occur in all four calibration sets. As shown in Table S1 in supplementary material, most of the nodes from Magallon *et al.*'s (2015) calibration sets are estimated to be older than those from Bell *et al.*'s (2010). Since fossils selected for calibration may not be the oldest members of the clade, and knowledge of intrafamilial phylogenetic relationship may be insufficient, node age estimates tend to be too young, thus highlighting Magallón *et al.*'s (2015) estimates as more conservative. Furthermore, considering the larger fossil record used for their age estimation, we recommend the application of Magallón *et al.*'s (2015) calibration sets to avoid inaccuracies in the computation of phylogenetic community structure and phylogenetic diversity.

Nevertheless, all four datasets are provided in the Supplementary Material for user's choice, thus permitting comparisons among the outcomes from the different calibration sets. Users who work with the entire euphyllophyte group should be sure to use 'ages_bell_exp_euphyllophyte.txt' (S10 in supplementary material), 'ages_bell_logn_euphyllophyte.txt' (S11 in supplementary material), 'ages_magallon_PL_euphyllophyte.txt' (S12 in supplementary material) or 'ages_magallon_UCLN_euphyllophyte.txt' (S13 in supplementary material), which include the age estimates for divergence times within and between monilophytes and gymnosperms as proposed by Hedges & Kumar (2009).

Pruning R20160415.new to the species lists from our case studies using the phylomatic function from Phylocom-4.2 was straightforward. However, the species *Pera glabrata* (Peraceae) from the FSN dataset, which could not be inserted into R20120829mod.new without changing its family affiliation to the Euphorbiaceae, is placed such that the tree topology suggested by APG IV (2016) was maintained. This was done because all species were classified into families recognized by APG IV. If one or more species had been classified among the families listed in Table 1, they would not be included in the community phylogeny by the phylomatic command unless the suffix '_NA' had been added to the name of its family in the 'species' file (not shown).

Because neither dataset contains *Amborella trichopoda*, the only extant representative of Amborellaceae, the angiosperm node in the community phylogeny is a singleton node that may impede the visualization of community phylogenies by some programs as well as its importation to the R environment. Therefore, we recommend the removal of this singleton node. This resulted in the bladj algorithm calibrating 47 (Bell *et al.* 2010) or 35 (Magallón *et al.* 2015) of 111 internal nodes in the FSN community tree; 21 calibrated nodes are the same in both calibration sets (Fig. 1). Forty (Bell *et al.* 2010) and 25 (Magallón *et al.* 2015) from 98 internal nodes were calibrated in the Eifel tree; from that, 21 are common ones. As previously outlined, Magallón *et al.* (2015) provide more age estimates for basal nodes, while Bell *et al.* (2010) also reported divergence times for more terminal nodes such as crown ages for the families of APG IV (2016). Thus, the nodes dated by Magallón *et al.* (2015) are concentrated on the left, basal, side of the phylogeny (Fig. 1), while Bell *et al.*'s (2010) age estimates are distributed more homogenously.

After calibration, the computation of measures of phylogenetic diversity and indexes of phylogenetic community structure were straightforward using the R environment or Phylocom 4.2. Although the 'ages' files show differences in mean age estimates (Tab. S1 in supplementary material), resulting in differences among phylogenetic trees (Fig. 1), outcomes from different sets of calibration points (i.e., exponential distribution (BEAST[a]) from Bell *et al.* (2010) and penalized likelihood from Magallón *et al.* (2015)) are significantly correlated (Fig. 2). Nonetheless, the correlation is not perfect; correlation coefficients ranging
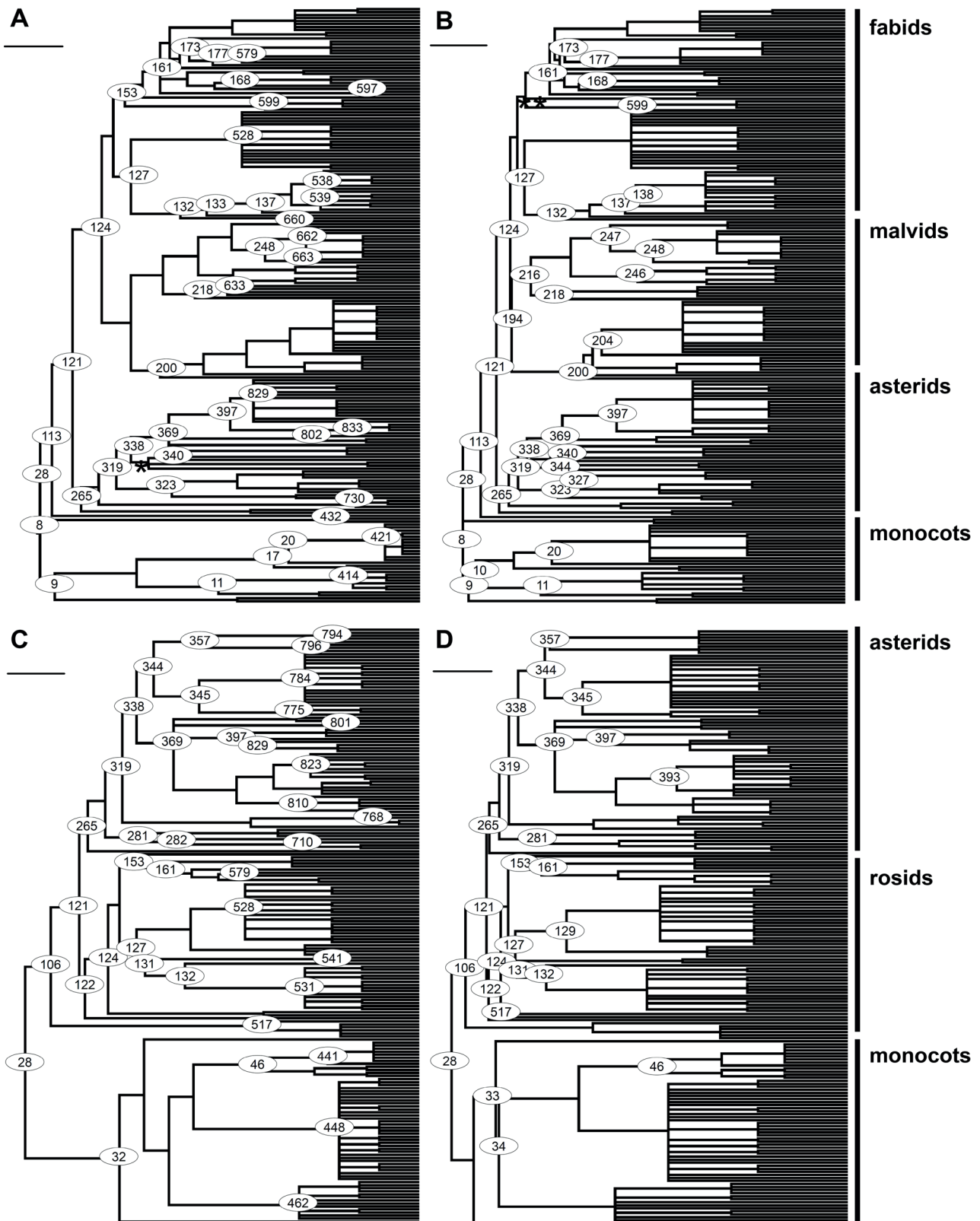
# Updated angiosperm family tree for analyzing phylogenetic diversity and community structure



**Figure 1.** Community trees for the Forest of Seu Nico Dynamics Plot (A, dated by exponential distribution from Bell *et al.* (2010); B, dated by penalized likelihood from Magallón *et al.* (2015)) and the Eifel Grassland dataset (C, dated by exponential distribution from Bell *et al.* (2010); D, dated by penalized likelihood from Magallón *et al.* (2015)) based on R20160415.new. Circles indicate calibrated internal nodes; numbers within circles refer to node IDs from Tab. S1 in supplementary material. * indicates nodes 339 (campanulids) and 344 (asterales_to_paracryphiales), ** indicates nodes 153 (COM) and 154 (oxalidales_to_malpighiales).

## A. Forest of Seu Nico Dynamics Plot
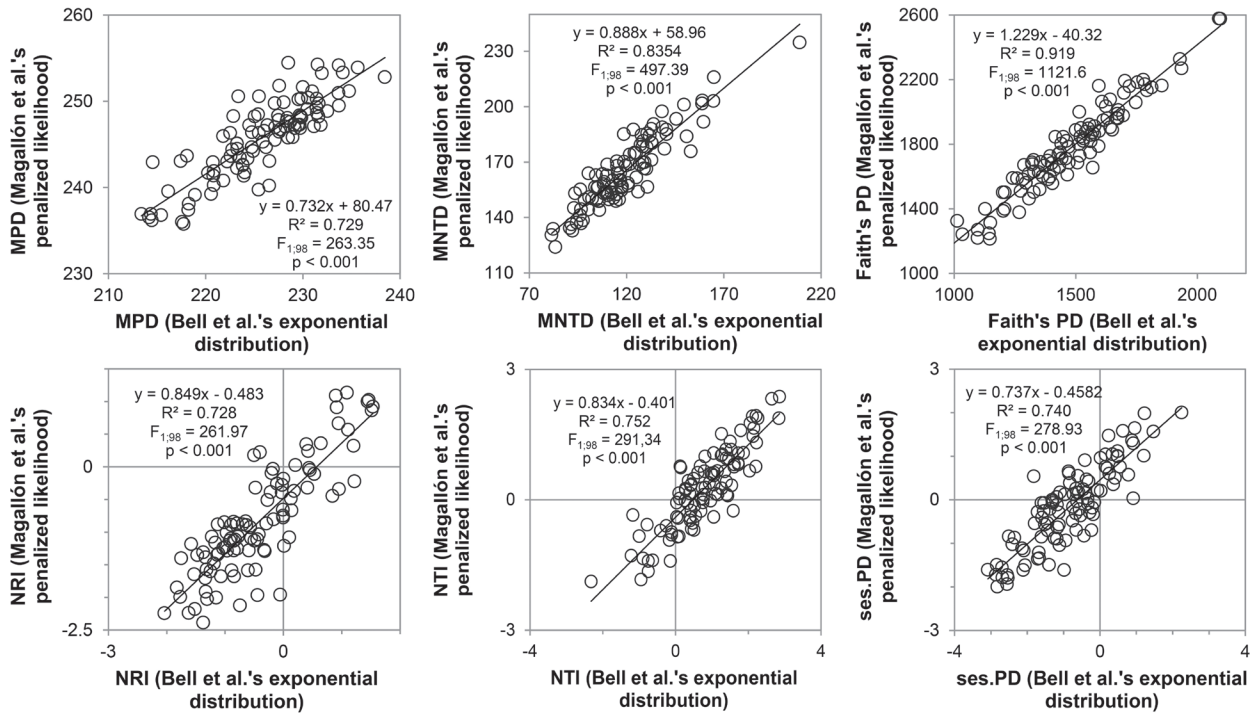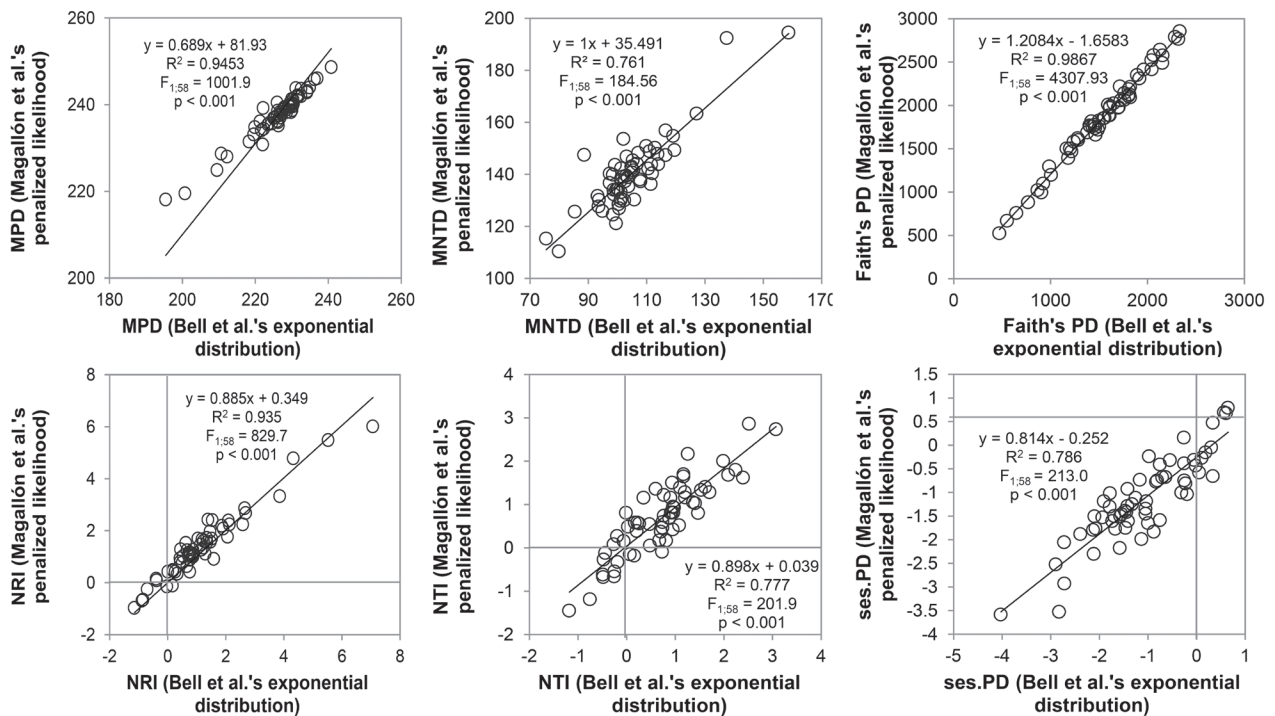


## B. Eifel Grassland Dataset



**Figure 2.** Correlations among the phylogenetic diversity (MPD is mean pairwise distance, MNTD is mean nearest neighbor distance, PD is phylogenetic diversity) and measures of phylogenetic community structure (NRI is net relatedness index, NTI is nearest taxon index and ses.PD is standard effect size of PF) calculated using R20160415.new and age estimates from Bell *et al.* (2010, exponential distribution) and Magallón *et al.* (2015, penalized likelihood).

from 0.7 to 0.9 indicate differences in the outcomes of the two calibration sets, which could certainly lead to ecological misinterpretation (Gastauer & Meira-Neto 2013). Furthermore, the finding that all measures, except Faith's PD, exhibited slopes less than one, indicates that outcomes computed using age estimates from Bell *et al*. (2010) tend to underestimate phylogenetic community structure, and especially NRI and NTI values. Ecological misinterpretation, as well as age underestimation, can certainly influence the interpretation of findings and the subsequent conclusions. To avoid this downfall, and to include evidence from as large a fossil record as possible in user-specific analyses and to reduce bias towards underestimating mean node age estimates, we recommend the application of the calibration sets from Magallón *et al*. (2015).

## Conclusion

Our goal was to provide an updated angiosperm phylogeny with updated minimum divergence times for the easy and straightforward computation of phylogenetic diversity and phylogenetic community structure. The phylogeny we present herein, R20160415.new, summarizes a recent review of angiosperm diversification (APG IV 2016) and makes these findings available for user-friendly computation of phylogenetic diversity and/or phylogenetic community structure. The inclusion of recently described angiosperm families and orders within R20160415.new justifies the relevance of this phylogeny in the analysis of phylogenetic community structure and phylogenetic diversity. Case studies have shown that using R20160415. new to analyze phylogenetic community structure or to compute phylogenetic diversity is straightforward. The chosen syntax of R20160415.new guides the user to insert species from the community of interest into the angiosperm family tree as indicated by APG IV (2016). The user receives feedback on unclear classifications, because invalid, yet still applied, family names without suffixes are not inserted by the phylomatic command, which allows the user to decide whether to refer to the actual (APG IV) classification or to an older one. We provide four different sets of node age estimates, from which a user can choose, but recommend the application of datasets excerpted from Magallón *et al*. (2015), as they are unbiased, do not erroneously underestimate age and represent a more extensive fossil record. We emphasize that R20160415.new is a hand-made supertree, and is not based on a proper phylogenetic analysis. Thus, calibration and dating may differ from biologically realistic divergence times. Nonetheless, in order to improve the precision of analyses, we recommend the consistent use of R20160415.new to ensure that up-to-date information about angiosperm evolution is included in the analysis of phylogenetic diversity and phylogenetic community structure, when more advanced techniques for phylogenetic reconstruction are not available.

## References

APG - Angiosperm Phylogeny Group III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Botanical Journal of the Linnean Society 161: 105-121.

APG - Angiosperm Phylogeny Group IV. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Botanical Journal of the Linnean Society 181: 1-20.

Arponen A, Zupan L. 2016. Representing hotspots of evolutionary history in systematic conservation planning for european mammals. In: Pellens R, Grandcolas P. (eds.) Biodiversity conservation and phylogenetic systematics. Genf, SpringerOpen. p. 265-285.

Bell CD, Soltis DE, Soltis OS. 2010. The age and diversification of the angiosperms re-revisited. American Journal of Botany 97: 1296-1303.

Boyle B, Hopkins N, Lu Z, *et al*. 2013. The taxonomic name resolution service: an online tool for automated standardization of plant names. BMC Bioinformatics 14: 16. doi: 10.1186/1471-2105-14-16

Cox CJ, Li B, Foster PG, Embley TM, Civáñ P. 2014. Conflicting phylogenies for early land plants are caused by composition biases among synonymous substitutions. Systematic Biology 63: 272 - 279.

Faith DP. 1992. Conservation evaluation and phylogenetic diversity. Biology Conservation 61: 1-10.

Filipowicz N, Renner SS. 2010. The worldwide holoparasitic Apodanthaceae confidently placed in the Cucurbitales by nuclear and mitochondrial gene trees. BMC Evolutionary Biology 10: 219. doi: 10.1186/1471-2148-10-219.

Fine PVA, Kembel SP. 2011. Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. Ecography 34: 552-565.

Forest F, Grenyer R, Rouget M, *et al*. 2007. Preserving the evolutionary potential of floras in biodiversity hotspots. Nature 445: 757-760.

Gastauer M, Leyh W, Meira-Neto JAA. 2015c. Tree diversity and dynamics of the Forest of Seu Nico. Biodiversity Data Journal 3: e5425. DOI: 10.3897/BDJ.3.e5425

Gastauer M, Meira-Neto JAA. 2013. Avoiding inaccuracies in tree calibration and phylogenetic community analysis using Phylocom 4.2. Ecological Informatics 15: 85-90.

Gastauer M, Meira-Neto JAA. 2014a. Interactions, environmental sorting and chance: Phylostructure of a Tropical Forest assembly. Folia Geobotanica 49: 443-459.

Gastauer M, Meira-Neto JAA. 2014b. Community dynamics in a species-rich old-growth forest patch from Viçosa, Minas Gerais, Southeastern Brazil. Acta Botanica Brasilica 27: 270-285.

Gastauer M, Meira-Neto JAA. 2016. An enhanced calibration of a recently released megatree for the analysis of phylogenetic diversity. Brazilian Journal of Biology 76. DOI: dx.doi.org/10.1590/1519-6984.20814.

Gastauer M, Saporetti-Júnior AW, Magnago LFS, Cavender-Bares J., Meira-Neto JAA. 2015a. The hypothesis of sympatric speciation as the dominant generator of endemism in a global hotspot of biodiversity. Ecology and Evolution: ece3.1761. doi: 10.1002/ece3.1761

Gastauer M, Sobral EG, Meira-Neto JAA. 2015b. Preservation of primary forest characteristics despite fragmentation and isolation in a forest remnant from Viçosa, MG, Brazil. Revista Árvore 39: 985-994.

Hedges SB, Kumar S. 2009. The timetree of life. New York, Oxford University Press.

Huang J, Huang J, Liu C, Zhang J, Lu X, Ma K. 2016. Diversity hotspots and conservation gaps for the Chinese endemic seed flora. Biological Conservation. doi:10.1007/s10531-015-1027-0.

Kembel SW, Cowan PD, Helmus MR, *et al*. 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463-1464.

Kembel SW, Hubbell SP. 2006. The phylogenetic structure of a neotropical forest tree community. Ecology 87: S86-S99.

Lamarre GPA, Amoretti DS, Baraloto B, Bénéluz F, Mesones I, Fine PFA. 2016. Phylogenetic overdispersion in Lepidoptera communities of Amazonian white-sand forests. Biotropica 48: 101-109.

Magallón S, Gómez-Acevedo S, Sánchez-Reyes LL, Hernández-Hernández T. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytologist 207: 437-453.

Miazaki AS, Gastauer M, Meira-Neto JAA. 2015. Environmental severity promotes phylogenetic clustering in campo rupestre vegetation. Acta Botanica Brasilica 29: 563-568.

Missouri Botanical Garden. 2016. http://www.tropicos.org. 12 Aug. 2016.

R Core Team. 2016. R: A language and environment for statistical computing. Vienna, R Foundation for Statistical Computing. https://www.R-project.org/.

Stevens PF. 2016. Angiosperm Phylogeny Website. http://www.mobot.org/MOBOT/research/APweb/. 12 Aug. 2016.

Webb CO, Ackerley DD, McPeek MA, Donoghue MJ. 2002. Phylogenies and community ecology. Annual Review of Ecology, Evolution, and Systematics 33: 475-505.

Webb CO, Donoghue M.J. 2005. Phylomatic: tree assembly for applied phylogenetics. Molecular Ecology Notes 5: 181-183.

Xiang QY, Thomas DT, Xiang QP. 2011. Resolving and dating the phylogeny of Cornales - effects of sampling, data partitions and fossil calibrations. Molecular Phylogenetics and Evolution 50: 123-138.