# Phylogenetic conservatism explains why plants are more likely to produce fleshy fruits in the tropics

Gang Wang[1], Anthony Ives[2], Hua Zhu[3], Yunhong Tan[1], Si-Chong Chen[4], Jie Yang[1], and Bo Wang[5]

[1]Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences
[2]UW-Madison
[3]Xishuangbanna Tropical Botanical Garden
[4]Royal Botanic Gardens Kew
[5]Anhui University

May 28, 2020

## Abstract

Plant functional traits often vary among species due to multiple factors. Here, using a dataset consisting of 9370 plant species from East Asia, we found that growth form and climate region explained only 1.7% and 0.3%, respectively, of the variance in fruit type, while phylogenetic conservatism explained 79.5%. Furthermore, phylogenetic conservatism was distributed from the base to the tips of the phylogeny, implying that fruit type reflects both ancient and recent phylogenetic relationships. When phylogeny was not considered, growth form and climate region explained 20.6% and 1.5% of the variation, implying that the association between fruit type and growth form is due primarily to phylogenetic conservatism of both traits, as opposed to correlated evolution. Our results highlight the dominant role of phylogenetic conservatism in explaining the pattern of fruit type can be revealed by parsing out the contributions of explanatory variables and phylogeny to the variance in species' traits.

## Introduction

The divergence of functional traits is usually influenced by multiple factors, and disentangling the relative contributions of these factors is important for understanding the evolutionary drivers of traits. Fruit type is a fundamental trait in plant biology, as it is a key component for seed dispersal, thereby linking reproduction of individual plants to regeneration of ecological communities (Howe & Smallwood 1982; Jordano *et al.* 2007; Seale & Nakayama 2020). Fruits can be dichotomously classified as fleshy or dry. The composition of fleshy versus dry fruits varies greatly across communities and climate regions, with an increasing proportion of fleshy-fruited species occurring with both decreasing latitude and altitude at a global scale (Willson *et al.* 1989; Almeida-Neto *et al.* 2008; Chen *et al.* 2017; Albert *et al.* 2018; Zhao *et al.* 2018). Although many factors have been proposed to explain this pattern, the relative importance of each factor has not yet been resolved. Here, we analyze a very large dataset to disentangle the relative contributions to the variation in fruit type of three factors that are most often included as possible explanations for fleshy versus dry fruits: plant growth form, environmental/climate factors (as summarized by the climate region in which species occur), and phylogenetic conservatism (evolutionary history).

Spatial variation in climate mediates the latitudinal and altitudinal patterns in functional traits of plants (Reich & Oleksyn 2004; Moles *et al.* 2007; Diaz *et al.* 2016). In general, precipitation is positively associated with the proportion of fleshy fruits (Correa *et al.* 2015; Chen *et al.* 2017; Zhao *et al.* 2018), as water availability is a crucial factor for the development of fleshy fruits (Coombe 1976). High temperature is also associated with an increasing proportion of species producing fleshy fruits (Willson *et al.* 1989; Zhao

1

*et al.* 2018), possibly because succulent fleshy fruits cannot survive freezing conditions (Burke *et al.* 1976). By analyzing the effect of 15 climatic variables on the latitudinal pattern in fruit types of 4008 Australian species, Chen *et al.* (2017) showed that wet, warm, and stable climates predicted a higher proportion of fleshy-fruited species. Furthermore, frugivory may be another possible explanation, as tropics usually hold more species and greater abundance of fruit-eating animals than temperate regions (Kissling *et al.* 2009; Onstein *et al.* 2017).

Plant growth form is often a predictor of other functional traits, for example, seed mass, stem specific density, and leaf area (Moles *et al.* 2005; Diaz *et al.* 2016). Fruit type is also strongly related to plant growth form, showing that trees and shrubs are more likely to produce fleshy fruits, while herbs are more likely to produce dry fruits (Chen *et al.* 2004; Bolmgren & Eriksson 2005; Cortes-Flores *et al.* 2013). de Queiroz (2002) proposed that the cost of producing fleshy fruits may reduce seed output, which would have proportionally more negative effects on small and short-lived plants. Furthermore, in the context of apparency theory (Feeny 1976), large and long-lived woody species that produce fleshy fruits may better advertise to frugivores than herbaceous fleshy species. Using a global dataset, Moles *et al.* (2009) showed a remarkable relationship between latitude and plant height, indicating that more trees occurred in the tropics than in the temperate regions. In this way, the latitudinal gradient in fruit type could be a reflection of the latitudinal gradient in growth form.

Phylogenetic signal of specific functional traits may be a consequence of either adaptation among phylogenetically related species that share similar environments or phylogenetic conservatism (Blomberg *et al.* 2003). Although the shift between fleshy and dry fruit occurred independently in many plant lineages as a response to environmental filters (Bolmgren & Eriksson 2005), a significant phylogenetic signal can still be detected (Herrera 2002; Chen *et al.* 2017), indicating that phylogenetic conservatism is a potential factor explaining fruit type. We therefore may expect the different composition of fruit types among floras to reflect their evolutionary histories. Furthermore, plant growth form may be also phylogenetically controlled, as are many other plant functional traits such as seed mass (Moles *et al.* 2005), leaf traits (Pearse & Hipp 2009) and root traits (Valverde-Barrantes *et al.* 2017). Therefore, we further expected that the phylogenetic conservatism might play an important role in determining the divergence of fruit type.

In summary, plant growth form, climatic factors, and phylogenetic conservatism may account, at least in part, for the pattern of fleshy versus dry fruit types (Bolmgren & Eriksson 2005; Cortes-Flores *et al.* 2013; Chen *et al.* 2017). However, disentangling the relative contributions of each factor is a challenge. Two common methods have been performed to partition variance of a target functional trait among possible predictor variables. Phylogenetic generalized least squares models (PGLS) with the phylogenetic tree treated as the error covariance structure can be used to compare the relative contributions among predictor variables after accounting for phylogenetic non-independence of species; however, PGLS models do not directly give a way to partition sources of variation explaining a trait. Nested ANOVAs and linear mixed models with several taxonomical levels (e.g., order, family, and genus) treated as nested random effects can be used to disentangle the relative contributions among multiple predictors, including phylogeny (Chen *et al.* 2004; Valverde-Barrantes *et al.* 2017). However, these models only consider the nested hierarchy of a few taxonomical levels, which oversimplifies the phylogeny of related taxa. Furthermore, interpreting the partition of contributions of multiple factors to the total explained variance ($R^2$) in the linear mixed model may be confounded when there is correlation between predictor variables (e.g., latitude and growth form). Therefore, to assess the roles of plant growth form, climate region, and phylogenetic conservatism on whether species produce fleshy or dry fruit, we used a phylogenetic partial $R^2$ that measures the reduction of explained variance when each factor is removed from a full model containing all other predictors (Ives & Li 2018; Ives 2019). Thus, the partial $R^2$ values give the amount of variation explained by a predictor by asking how much of the explained variation is lost when the factor is removed.

In this study, we give a quantitative analysis of the relative contributions of plant growth form, climate region, and phylogenetic conservatism to the variation of fruit type with a data set including 9,370 species from Southwest China which represent a tropical, subtropical, and temperate flora. We predicted that (i)

woody species produce more fleshy fruits than herbs; (ii) more fleshy-fruit species occur in tropical than in subtropical and temperate floras climate regions; and (iii) fruit type is phylogenetically conserved. Because we found that phylogenetic conservativism in fruit type was extremely strong, we further investigated whether this was due to ancient or recent phylogenetic relationships towards the base or tips of the phylogeny, respectively. Finally, we investigated the phylogenetic conservatism of climate region and plant growth form: if fruit type were to show greater phylogenetic conservatism than either of these predictor variables, then it would not be surprising that they are poor predictors, because the greater phylogenetic conservatism of fruit type would limit its evolutionary response to the predictors.

**Material and methods**

Study area

This study focused on three climate regions in Yunnan Province, southwest China (21°09'–29deg15'N and 97deg32'–106deg12' E). The three regions are similar in area, but differ greatly in climate, geography, and vegetation (Table 1). The southern region has a tropical monsoon climate and contains a tropical flora that was shaped by the strike-slip extrusion of Indochina plate. The northwestern region has a temperate climate and contains a temperate flora shaped by the uplift of the Himalaya-Tibetan Plateau. The central region has a subtropical climate and contains a subtropical flora with a combination of nearly equal proportions of species from both northwestern and southern floras (Zhu 2012; Liu *et al.* 2017). The overall floras of the regions combined are representative of East Asia (Zhu 2012; Liu *et al.* 2017).

Floristic composition

A comprehensive list of the native seed plants of the above floras was obtained from several floristic inventories (Peng 1998; Zhu & Yan 2009, 2012; Zhu 2015), the Flora of Yunnan (Wu 1977-2006), and the database of seed plants from herbarium of Kunming Institute of Botany, Chinese Academy of Sciences. The nomenclature of these species was checked and standardized according to The Plant List version 1.1 (TPL, available at http://www.theplantlist.org). All the intraspecific taxa (e.g., subspecies, variety and forma) and introduced species were excluded. In total, 9,370 seed plant species were included in this study, belonging to 1,860 genera and 212 families. Details about the data collection are described in Zhu (2012) and Liu *et al.* (2017), both of which discussed the biogeographical divergences among these three floras.

Functional traits

Fruit type (fleshy vs. dry fruit) and plant growth form (woody vs. herbaceous) data were collected from published floras: *Flora Repubulicae Popularis Sinicae* (Flora Repubulicae Popularis Sinicae Editorial Committee 1959-2004), the *Flora of Yunnan* (Wu 1977-2006), and *Iconographia cormophytorum sinicorum* (Institute of Botany, 1972-2002). For some species, there was no information in the published floras, so their fruit type and growth form were based on *Chinese Virtual Herbarium* (http://www.cvh.ac.cn/) and *Plant Photo Bank of China* (http://www.plantphoto.cn/). Fruits (sometimes dispersal diaspores) were classified as fleshy if there were apparent fleshy pericarps or appendages when mature; otherwise, they were classified as dry (Willson *et al.* 1989; Chen *et al.* 2017).

Phylogeny tree construction

A phylogenetic tree including all the 9370 species was obtained following Liu *et al.* (2017), which used the same species matrix. The phylogeny was constructed with the R package "S.PhyloMaker" (Qian & Jin 2016), with a species list as primary input data. An updated version of the supertree in Zanne *et al.* (2014), including 30,771 species of all families of extant seed plants based on the available genes in GenBank, was used as the backbone in the package "S.PhyloMaker". Genera and species that were not included in the backbone tree were added as polytomies within their families or genera.

Statistical analyses

The three regions differ greatly in climate, topography, and altitude range (Table 1). Furthermore, which regions are occupied by a species is the outcome of many factors. Therefore, in the statistical analyses we

3

treat region as a synoptic trait that encompasses many unspecified environmental factors. The three regions are contiguous, and we assume that, over the evolutionary time scale represented by the diversity of the 9,370 species in the data set, there is little constraint caused by dispersal; this assumption is confirmed by the analyses (see "Results").

To partition the relative contributions of climate region, plant growth form, and phylogenetic conservatism on the variation in fruit type, we used the R package "rr2" (Ives & Li 2018; Ives 2019) to calculate partial $R^2$s for phylogenetic logistic regression models fit using the *phyloglm* function in the R package "phylolm" (Tung Ho & Ane 2014). The *phyloglm* function allows for the calculation of only one type of partial $R^2$ in the rr2 package, $R^2_{lik}$, that is based on the likelihood. $R^2_{lik}$is the appropriate $R^2$ when comparing models according to the statistical significance of differences between them. The partial $R^2_{lik}$ for each factor was calculated by comparing the full model with reduced models in which a given factor was removed, and measuring the consequent reduction in explained variance. The full model is a logistic regression model with fruit type (fleshy and dry) as the dependent variable, climate region (tropic, subtropical, and temperate) and growth form (woody and herbaceous) as independent variables, and phylogeny as covariances in the residual variation.

To give more information about the pattern of phylogenetic conservatism, we performed the analyses not only with the full time-scaled phylogeny, but also a phylogeny in which relatively recent phylogenetic diversification was removed. Specifically, we created phylogenies with reduced "recent" phylogenetic structure by collapsing nodes above a given threshold together to form a "star". For example, for a threshold of 0.67, any node above the 66.7% (2/3) mark on the phylogeny was collapsed so that evolution above this threshold was assumed to occur independently among species (Fig. 1). The specific threshold of 0.67 corresponds roughly to the taxonomic scale of families. To determine the importance phylogenetic patterns across the entire depth of the phylogeny, we performed the same analysis using thresholds from 1 down to 0.2. Re-analyzing the data across this range of thresholds addresses the relative importance of "recent" versus "ancient" species relationships in explaining phylogenetic conservatism.

The logistic regression showed that fruit type was poorly explained by climate region and growth form, but the residual variation was well-explained by the phylogeny. Because we had expected climate region and growth form to be good predictors of fruit type, we investigated all three variables – fruit type, climate region, and growth form – separately to determine whether they all show similar patterns of phylogenetic conservatism. If climate region and/or growth form show less phylogenetic conservatism than fruit type, this would suggest that climate region and/or growth form are phylogenetically more labile and therefore might not be expected to predict fruit type. If climate region and/or growth form show the same phylogenetic conservatism as fruit type, then this, in combination with the low explanatory power they have of fruit type, would imply that their evolution is uncorrelated to fruit type.

### Results

Of the 9,370 species (combined dataset of the three floras), 2,635 species (28.1%) produced fleshy fruits. Temperate flora contained a lower proportion of fleshy-fruited species (22.5%) than both tropical flora (38.7%) and subtropical floras (32.6%; Fig. 2a, Table S1 in Supporting Information). Woody species contained a higher proportion of fleshy-fruited species (55.6%) than herbaceous species (9.6%) (Fig. 2b, Table S1).

Phylogenetic conservatism explained the vast majority of fruit type variation among floras (partial $R^2_{lik}$ = 79.5%), while climate region and growth form explained little variation, with partial $R^2_{lik}$ values of 0.3% and 1.7%, respectively (Table 2, Fig. 2c, d). To investigate the depth at which this conservatism occurred within the phylogeny, we repeated the logistic regression using a phylogeny in which nodes above the 0.67 threshold from the base of the phylogeny were collapsed (Fig. 1b). This threshold corresponds roughly to the taxonomic level of families, so this phylogeny removed phylogenetic information within families. The phylogenetic conservatism occurred both towards the tips and towards the base of the phylogenetic tree: when using the 0.67-threshold phylogeny, the partial $R^2_{lik}$ for phylogeny was 41.6%, implying that the phylogenetic information towards the base of the phylogeny could explain roughly half as much variation

as the full phylogeny. Furthermore, the partial $R^2_{lik}$ values for plant growth form and climate region were 25.1% and 3.9% (Table 2), implying that plant growth explains more of the distribution of fruit type among species when the explanatory information provided towards the tips of the phylogeny is not available.

The low partial $R^2_{lik}$ for growth form can be better understood by considering the regression of growth form on fruit type without the phylogeny or climate regions. In the model with growth form alone (full model GF, Table 2), the $R^2_{lik}$ is 22.8% implying that growth form explains a moderate amount of variation in fruit type. This is expected, because 55.6% of the wood species had fleshy fruit, compared to only 9.6% of the herbaceous species (Fig. 2b, Table S1). However, in the model with growth form and phylogeny (full model GF+PHY, Table 2), the partial $R^2_{lik}$ for growth form is 1.5%. Therefore, when taking account of phylogeny, growth form gives little additional information about fruit type. This suggests that the apparent association between fruit type and growth form (Spearman rank correlation = 0.503, Table S1, $R^2_{lik}$ = 22.8%, Table 2) observed in this dataset cannot be distinguished from a spurious effect of both fruit type and growth form being strongly conserved phylogenetically. In contrast to growth form, the $R^2_{lik}$ for climate region remains low even in the absence of phylogeny ($R^2_{lik}$ = 4.2%, full model CR, Table 2). This implies that phylogeny does not play a role in the poor explanatory power of climate regions on fruit type.

To investigate lack of explanatory power of growth form and climate region on fruit type in more detail, we investigated the phylogenetic conservatism for each trait separately. Specifically, we performed analyses with each trait as the dependent variable and no independent variables. Both fruit type and plant growth form had very strong phylogenetic signal (low values of α of the Ornstein-Uhlenbeck model of evolution employed by phyloglm, Tung Ho and Ané 2014) and $R^2_{lik}$ values of 86.3% and 86.8%, respectively, when the full phylogeny used (Table 3). When the 0.67-threshold phylogeny was used for the analyses, the $R^2_{lik}$ values decreased to 46.8% and 53.2%, respectively. This implies that phylogenetic conservatism occurred both towards the base of the phylogeny (because the traits $R^2_{lik}$ values were high) and also towards the tip (because the traits $R^2_{lik}$ values were not as high as for the full phylogeny). The results for fruit type and plant growth form contrast those for climate regions. For tropical, subtropical, and temperate climate regions, the $R^2_{lik}$ values were 20.0%, 0%, and 12.2%, respectively. Furthermore, the $R^2_{lik}$ values of the three climate regions computed for the 0.67-threshold phylogeny were similar to those for the full phylogeny, implying that most of the information available from the phylogeny occurred towards the base of the tree.

To obtain a more-complete picture of the pattern of phylogenetic conservatism of the three traits, for each trait we computed the $R^2_{lik}$ across a range of threshold phylogenies, from 1 (full phylogeny) down to 0.2 (only 20% phylogenetic information near the basal was maintained) (Fig. 3). Both fruit type and growth form show very strong phylogenetic conservatism, and as the threshold above which nodes are collapsed increases, the values of $R^2_{lik}$ increase roughly linearly. This implies that the phylogenetic information underlying the values of $R^2_{lik}$ are even distributed across the phylogeny. In contrast, the values of $R^2_{lik}$ for tropical and temperate climate regions stay roughly the same until the threshold drops below 0.8, implying that the phylogenetic relationships above this threshold provide little information about whether species occur in the two climate regions. The values of $R^2_{lik}$ only drop substantially when they are below the 0.8 threshold, implying that phylogenetic conservatism occurs towards the base of the phylogeny. For the subtropical climate region, $R^2_{lik}$ is always close to zero, implying that whether or not a species occurs in the subtropical climate region is not explained by phylogeny.

**Discussion**

Across a broad taxonomic and geographic range of plant species in SW China, plant growth form and climate region had little power to explain whether a plant species had fleshy or dry fruit. In contrast, fruit type exhibited strong phylogenetic conservatism. This conservatism was distributed over the entire phylogeny; in other words, broad phylogenetic relationships established at the base of the phylogeny, as well as recent phylogenetic relationships towards the tips of the phylogeny, all contributed to the strong phylogenetic conservatism of fruit type.

Fleshy fruit pulp has been hypothesized to evolve as a defense against seed predation (Mack 2000), and fleshy

5

fruit are secondarily associated with endozoochory to promote seed dispersal (Howe & Smallwood 1982; Onstein *et al.* 2017). Therefore, weak phylogenetic conservatism of fruit type might be expected because of the strong selection from seed predators and frugivore seed dispersers. In contrast to this expectation, a strong phylogenetic signal of fruit type was detected in this study which is consistent with some previous studies that focused at different scales of both geography and taxa (Herrera 2002; Chen*et al.* 2017). Fleming and Kress (2011) reported that the first appearance of the core angiosperm families producing fleshy fruits were earlier than the first appearance of the families of their major fruit eating animals. Therefore, the phylogeny signal residing towards the base of the phylogeny might be independent of coevolution with frugivores, and instead the later-appearing frugivores evolved to take advantage of the preexisting fruit types. Other fruit/seed traits, such as seed size, fruit size and fruit color appear to evolve predominately under the selection pressure of fruit-eating animals; these traits also show phylogenetic signal, which implies that selection may generate phylogenetic signal if related species are under the same selective pressures (Chen *et al.* 2004; Moles *et al.* 2005). In our study, the majority of families (67.8%) exclusively produce fleshy or dry fruits, with shifts between dry and fleshy fruit occurring in the remaining families.

We had expected both plant growth form and climate region to explain much of the variation in fruit type. Our analyses suggest that the reason they do not provide much explanation differs between the two factors. Like fruit type, growth form itself shows strong phylogenetic conservatism that is distributed over the phylogeny (Fig. 3). When phylogeny is not included, growth form explains 22.8% of the variation in fruit type, while this drops to 1.5% when phylogeny is included (Table 2). As another useful comparison, the partial $R^2_{lik}$ for growth form is 1.7% when the full phylogeny is used, but this increases to 25.1% when the 0.67-threshold phylogeny is used which removes the phylogeny information in the top 1/3 of the phylogeny. These two patterns suggest that fruit type and growth form are evolving largely independently, and their apparent association in the raw data – with 55.6% of the wood species and only 9.5% of the herbaceous species having fleshy fruit – is due largely to the phylogenetic conservatism of both traits.

In contrast to growth form, the climate region in which species occur does not show strong conservatism. What conservatism tropical and temperate climate regions have occurs towards the base of the phylogeny, implying that there are major clades being associated with tropical and temperate regions, but at finer scales towards the tips of the phylogeny, the role of phylogeny diminishes (Fig. 3). Thus, the explanation for climate region having little effect on fruit type is that climate region is relatively phylogenetically labile, changing up the phylogeny more rapidly than fruit type, especially towards the tips.

In contrast to our study, previous studies have shown effects of climate region (or environmental factors in general) on fruit type (Willson*et al.* 1989; Bolmgren & Eriksson 2005; Chen *et al.* 2017; Zhao *et al.* 2018). There are several possible explanations for our differing results. We confined our study to plants occurring in SW China, a region known for its biogeographic and topographic/environmental diversity (Zhu 2012). Because tropical, subtropical, and temperate climate regions are in close proximity, we suspect that dispersal of species among regions is rapid on the phylogenetic time scales represented by our study plants. This suggests that dispersal and local speciation may be responsible for the lack of phylogenetic conservatism of those traits that determine the climate region in which species live. Thus, the traits that determine success in a climate region are evolutionarily labile, and the trait of fruit type plays at most a minor role in determining success in climate regions.

There may also be methodological explanations for why our results differ from many previous studies, that have investigated phylogenetic patterns for traits using different taxonomic levels, for example in nested ANOVAs (Chen *et al.* 2004; Valverde-Barrantes *et al.* 2017). Our approach allows a finer-scale assessment of the role of phylogeny. For example, using the 0.67-threshold phylogeny, in which phylogenetic information above roughly the family level is removed, growth form was a moderately strong predictor of fruit type, explaining 25.1% of the variance. However, using the full phylogeny, this dropped to 1.7%. Therefore, if nested ANOVAs did not account for the fine relationships towards the tips of the phylogeny, then they would miss the full effects of phylogenetic conservatism.

Exploring the drivers of the geographic variation in plant functional traits, especially along latitude or/and

altitude gradients, has long attracted attention in eco-biogeography (Reich & Oleksyn 2004; Moles *et al.* 2007; Moles *et al.* 2009; Diaz *et al.* 2016; Valverde-Barrantes *et al.* 2017). Several models have been introduced to account for phylogenetic non-independence among species, although they have not quantified the contribution of phylogeny itself to trait variation (Swenson & Enquist 2007; Chen *et al.* 2017). However, quantifying the effect of phylogeny in comparison to other factors is useful (Valverde-Barrantes *et al.* 2017; this study). Partial $R^2$s can disentangle the relative contributions of both predictor variables (climate region and growth form in this study) and covariances (phylogeny and other possible random effects). Combined with the findings from recent empirical studies on other important plant functional traits (Moles *et al.* 2005; Swenson & Enquist 2007; Valverde-Barrantes *et al.* 2017), our results suggest that methods using the full phylogeny rather than earlier methods such as nested ANOVAs have more power to explore the phylogenetic patterns of functional traits.

## Acknowledgements

## References

Albert, S., Flores, O., Rouget, M., Wilding, N. & Strasberg, D. (2018). Why are woody plants fleshy-fruited at low elevations? Evidence from a high-elevation oceanic island. *J. Veg. Sci.* , 29, 847-858.

Almeida-Neto, M., Campassi, F., Galetti, M., Jordano, P. & Oliveira-Filho, A. (2008). Vertebrate dispersal syndromes along the Atlantic forest: broad-scale patterns and macroecological correlates. *Global Ecol. Biogeogr.* , 17, 503-513.

Blomberg, S.P., Garland, T. & Ives, A.R. (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution* , 57, 717-745.

Bolmgren, K. & Eriksson, O. (2005). Fleshy fruits - origins, niche shifts, and diversification. *Oikos* , 109, 255-272.

Burke, M.J., Gusta, L.V., Quamme, H.A., Weiser, C.J. & Li, P.H. (1976). Freezing and injury in plants. *Ann. Rev. Plant Physiol.* , 27, 507-528.

Chen, J., Fleming, T.H., Zhang, L., Wang, H. & Liu, Y. (2004). Patterns of fruit traits in a tropical rainforest in Xishuangbanna, SW China. *Acta. Oecol.* , 26, 157-164.

Chen, S.-C., Cornwell, W.K., Zhang, H.-X. & Moles, A.T. (2017). Plants show more flesh in the tropics: variation in fruit type along latitudinal and climatic gradients. *Ecography* , 40, 531-538.

Coombe, B.G. (1976). The development of fleshy fruits. *Ann. Rev. Plant Physiol.* , 27, 207-228.

Correa, D.F., Alvarez, E. & Stevenson, P.R. (2015). Plant dispersal systems in Neotropical forests: availability of dispersal agents or availability of resources for constructing zoochorous fruits? *Global Ecol. Biogeogr.* , 24, 203-214.

Cortes-Flores, J., Andresen, E., Cornejo-Tenorio, G. & Ibarra-Manriquez, G. (2013). Fruiting phenology of seed dispersal syndromes in a Mexican Neotropical temperate forest. *For. Ecol. Manage.* , 289, 445-454.

de Queiroz, A. (2002). Contingent predictability in evolution: key traits and diversification. *Syst. Biol.* , 51, 917-929.

Diaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S. *et al.* (2016). The global spectrum of plant form and function. *Nature* , 529, 167-171.

Feeny, P. (1976). Plant apparency and chemical defense. In: *Recent Advances in Phytochemistry* (eds. Wallace, JW & Mansell, RL). Plenum Press, New York, pp. 1-40.

Fleming, T.H. & Kress, W.J. (2011). A brief history of fruits and frugivores. *Acta. Oecol.* , 37, 521-530.

Flora Repubulicae Popularis Sinicae Editorial Committee (1959-2004). *Flora Repubulicae Popularis Sinicae* . Science Press (In Chinese), Beijing.

Herrera, C.M. (2002). Seed dispersal by vertebrates. In: *Plant-Animal Interactions: an Evolutionary Approach* (eds. Herrera, CM & Pellmyr, O). Blackwell Science Oxford, UK, pp. 185-210.

Howe, H.F. & Smallwood, J. (1982). Ecology of seed dispersal. *Annu. Rev. Ecol. Syst.* , 13, 201-228.

Institute of Botany the Chinese Academy of Sciences (1972-2002). *Iconographia Cormophytorum Sinicorum* . Science Press (In Chinese), Beijing.

Ives, A.R. (2019). $R^2$s for correlated data: phylogenetic models, LMMs, and GLMMs. *Syst. Biol.* , 68, 234-251.

Ives, A.R. & Li, D. (2018). rr2: An R package to calculate $R^2$s for regression models. *Journal of Open Source Software* , 3, 1028.

Jordano, P., Garcia, C., Godoy, J.A. & Garcia-Castano, J.L. (2007). Differential contribution of frugivores to complex seed dispersal patterns. *PNAS* , 104, 3278-3282.

Kissling, W.D., Bohning–Gaese, K. & Jetz, W. (2009). The global distribution of frugivory in birds. *Global Ecol. Biogeogr.* , 18, 150-162.

Liu, S., Zhu, H. & Yang, J. (2017). A phylogenetic perspective on biogeographical divergence of the flora in Yunnan, southwestern China. *Sci. Rep.* , 7, 43032.

Mack, A.L. (2000). Did fleshy fruit pulp evolve as a defence against seed loss rather than as a dispersal mechanism? *J. Biosciences* , 25, 93-97.

Moles, A.T., Ackerly, D.D., Tweddle, J.C., Dickie, J.B., Smith, R., Leishman, M.R. *et al.* (2007). Global patterns in seed size. *Global Ecol. Biogeogr.* , 16, 109-116.

Moles, A.T., Ackerly, D.D., Webb, C.O., Tweddle, J.C., Dickie, J.B. & Westoby, M. (2005). A brief history of seed size. *Science* , 307, 576-580.

Moles, A.T., Warton, D.I., Warman, L., Swenson, N.G., Laffan, S.W., Zanne, A.E. *et al.* (2009). Global patterns in plant height. *J. Ecol.* , 97, 923-932.

Onstein, R.E., Baker, W.J., Couvreur, T.L.P., Faurby, S., Svenning, J.-C. & Kissling, W.D. (2017). Frugivory-related traits promote speciation of tropical palms. *Nat. Ecol. Evol.* , 1, 1903-1911.

Pearse, I.S. & Hipp, A.L. (2009). Phylogenetic and trait similarity to a native species predict herbivory on non-native oaks. *PNAS* , 106, 18097-18102.

Peng, H. (1998). *The Seed Plants From Mt. Wuliang in South-Central Yunnan, China* . Yunnan Science and Technology Press (in Chinese), Kunming.

Qian, H. & Jin, Y. (2016). An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *J. Plant Ecol.* , 9, 233-239.

Reich, P.B. & Oleksyn, J. (2004). Global patterns of plant leaf N and P in relation to temperature and latitude. *PNAS* , 101, 11001-11006.

Seale, M. & Nakayama, N. (2020). From passive to informed: mechanical mechanisms of seed dispersal. *New Phytol.* , 225, 653-658.

Swenson, N.G. & Enquist, B.J. (2007). Ecological and evolutionary determinants of a key plant functional trait: Wood density and its community-wide variation across latitude and elevation. *Am. J. Bot.* , 94, 451-459.

Tung Ho, L.s. & Ane, C. (2014). A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol.* , 63, 397-408.

Valverde-Barrantes, O.J., Freschet, G.T., Roumet, C. & Blackwood, C.B. (2017). A worldview of root traits: the influence of ancestry, growth form, climate and mycorrhizal association on the functional trait variation of fine-root tissues in seed plants. *New Phytol.* , 215, 1562-1573.

Willson, M.F., Irvine, A.K. & Walsh, N.G. (1989). Vertebrate dispersal syndromes in some Australian and New-Zealand plant-communities, with geographic comparisons.*Biotropica* , 21, 133-147.

Wu, Z.Y. (1977-2006). *Flora Yunnanica* . Science Press (In Chinese), Beijing.

Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A., FitzJohn, R.G. *et al.* (2014). Three keys to the radiation of angiosperms into freezing environments.*Nature* , 506, 89-92.

Zhao, Y., Cao, H., Xu, W., Chen, G., Lian, J., Du, Y. *et al.* (2018). Contributions of precipitation and temperature to the large scale geographic distribution of fleshy-fruited plant species: growth form matters. *Sci. Rep.* , 8, 17017.

Zhu, H. (2012). Biogeographical divergence of the flora of Yunnan, southwestern China initiated by the uplift of Himalaya and extrusion of Indochina block. *Plos One* , 7, e45601.

Zhu, H. (2015). Biogeography of Shangri-la flora in southwestern China. *Phytotaxa* , 203, 231-244.

Zhu, H. & Yan, L.C. (2009).*List of Seed Plants in the Ailao Mts. of Yunnan Province, China* . Yunnan Science and Technology Press (in Chinese), Kunming.

Zhu, H. & Yan, L.C. (2012).*Native Seed Plants in Xishuangbanna of Yunnan* . Science Press (in Chinese), Beijing.

**Figure Legends**

**Figure 1** Phylogenies of 500 randomly selected species from the dataset of 9370 species. (a) 1-threshold, the full time scale phylogeny; (b) 0.67-threshold, the same phylogeny but when nodes above the 0.67 threshold (i.e., nodes in the top 1/3 of the phylogeny) are collapsed to remove the "recent" phylogenetic relationships; and (c) 0.40-threshold, phylogeny with nodes above the 0.40 threshold are collapsed. The 0.67 threshold corresponds roughly to taxonomic families. The full dataset of 9370 species was subsampled for better visualization.

**Figure 2** Distribution of fleshy-fruited versus dry-fruited species in different (a) climate regions, (b) growth forms and (c) phylogeny, and (d) the contribution of different factors to the variation in fruit type. Numbers in (a) and (b) are number of species for each fruit type.

**Figure 3** The values of $R^2_{lik}$ for phylogenetic models of each trait separately as "recent" phylogenetic relationships are removed from the full phylogeny. "Threshold" is the level in the full phylogeny above which nodes are collapsed; see Fig. 1 The age for each threshold was shown in the parentheses.

**Table 1** Summary of climate, vegetation and geographic information of the three study regions.

| Study Region | Southern Yunnan | Central Yunnan | Northwestern Yunnan |
| --- | --- | --- | --- |
| Area | 19690 km$^2$ | 25424 km$^2$ | 23870 km$^2$ |
| Latitude | 21°09' - 22deg36'N | 23°53' - 25deg11' N | 27°10' - 28deg27' N |
| Longitude | 99°58' - 101deg50' E | 100°32' - 101deg58' E | 98°53' - 99deg42' E |
| Altitude | 480 - 2430 m | 422 - 3157 m (mostly 1300 - 2200 m) | 1900 - 6740 m |
| Topography | low mountain-basin | middle mountain-valley | alpine-deep valley |
| Climate | tropical monsoon | Subtropical | temperate |

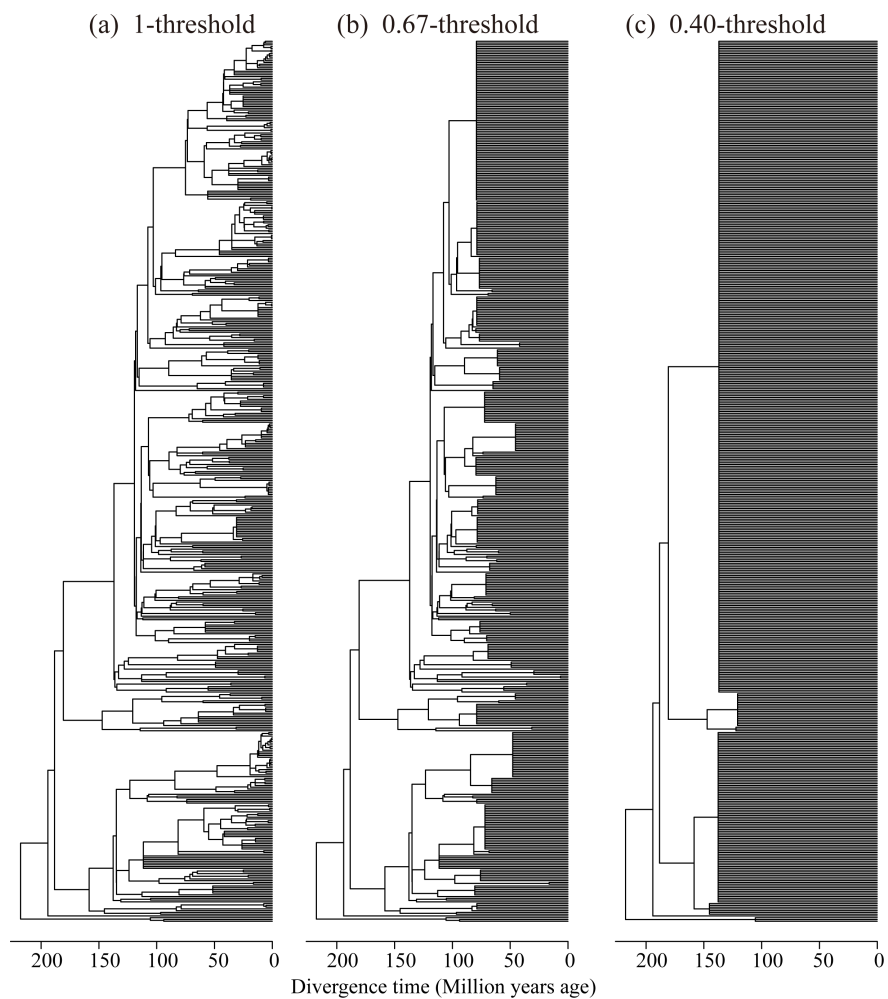| Study Region | Southern Yunnan | Central Yunnan | Northwestern Yunnan |
|---|---|---|---|
| Annual Mean Temperature | 18.9 - 22.6 | 14.9 - 18.9 | 4.7 - 16.5 |
| Annual Rainfall | 1130 - 1510 mm | 820 - 1090 mm | 620 - 940 mm |
| Dominate Vegetation | tropical rain forest | subtropical evergreen broad-leaf forests secondary *Pinus yunnanensis* forests | temperate sclerophyllous oak forests cold temperate coniferous forests |
| Number of Species | 3785 | 2972 | 6106 |

**Table 2** Partial contribution of each factor (growth form, GF; climate region, CR; and phylogeny, PHY) to the variation in fruit type for the full phylogeny and the 0.67-threshold phylogeny in which nodes in the top 1/3 of the phylogeny were collapsed. The partial contribution of a factor to a model is obtained by computing the partial $R^2_{lik}$ between the full model and the reduce model in which the factor is removed. The $\Delta$logLik is the change of log-likelihood between full and reduced models.

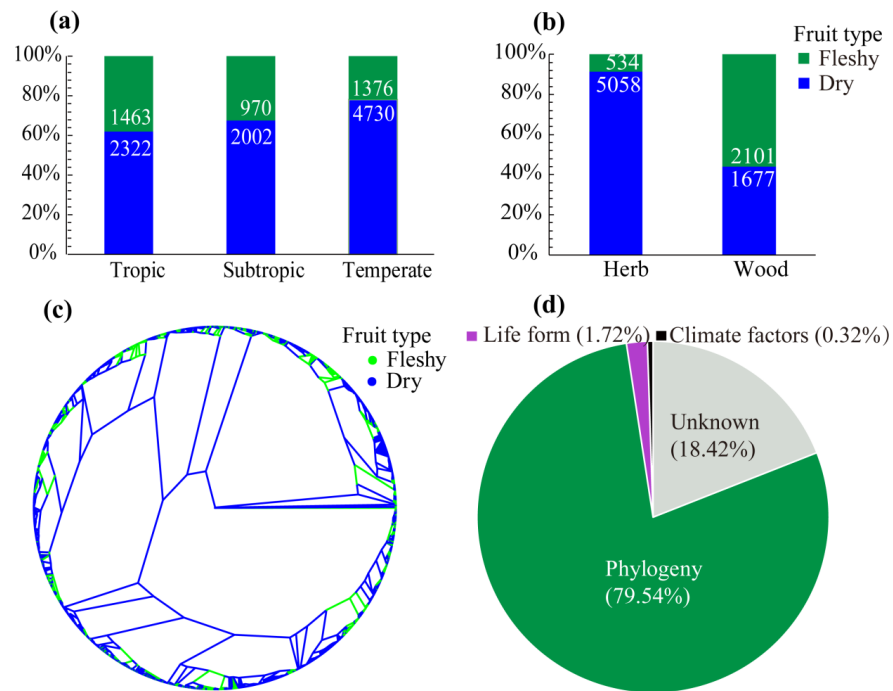| full model | reduced model | full phylogeny | full phylogeny | full phylogeny | 0.67-threshold phylogeny | 0.67-thresh |
|---|---|---|---|---|---|---|
| | | $R^2_{lik}$ | $\Delta$logLik | P | $R^2_{lik}$ | $\Delta$logLik |
| GF+CR+PHY | 1 | 0.866 | 4316.50 | <0.001 | 0.617 | 2625.39 |
| GF+CR+PHY | CZ+PHY | 0.017 | 19.11 | <0.001 | 0.251 | 680.61 |
| GF+CR+PHY | GF+PHY | 0.003 | 3.55 | 0.069 | 0.039 | 87.69 |
| GF+CR+PHY | GF+CZ | 0.795 | 3033.43 | <0.001 | 0.416 | 1342.32 |
| GF+PHY | GF | 0.800 | 3102.97 | <0.001 | 0.408 | 1327.72 |
| GF+PHY | PHY | 0.015 | 16.95 | <0.001 | 0.251 | 695.06 |
| CR+PHY | CZ | 0.855 | 4095.61 | <0.001 | 0.456 | 1743.01 |
| CR+PHY | PHY | 0.001 | 1.38 | 0.431 | 0.039 | 102.15 |
| GF+CR | GF | 0.015 | 73.09 | <0.001 | 0.015 | 73.09 |
| GF+CR | CZ | 0.206 | 1081.30 | <0.001 | 0.206 | 1081.30 |
| GF | 1 | 0.228 | 1209.98 | <0.001 | 0.228 | 1209.98 |
| CR | 1 | 0.042 | 201.78 | <0.001 | 0.042 | 201.78 |
| PHY | 1 | 0.863 | 4296.01 | <0.001 | 0.468 | 1842.64 |

**Table 3** Analyses of fruit type, growth form, and climate regions treated as dependent variables with no independent variables using logistic regression (phyloglm, Ho and Ane 2014). Values of α is the parameter from the Ornstein-Uhlenbeck model of evolution, and lower values correspond to stronger phylogenetic signal. The analyses were performed with both the full phylogeny and the 0.67-threshold phylogeny in which nodes in the top 1/3 of the phylogeny were collapsed.

| | Full phylogeny | Full phylogeny | 0.67-threshold phylogeny | 0.67-threshold phylogeny |
|---|---|---|---|---|
| | α | $R^2_{lik}$ | α | $R^2_{lik}$ |
| Fruit type | 0.779 | 0.863 | 1.507 | 0.468 |
| Growth form | 0.747 | 0.868 | 1.322 | 0.532 |
| Tropical | 7.652 | 0.200 | 3.322 | 0.183 |
| Subtropical | 54.330 | 0.000 | 8.434 | 0.009 |
| Temperate | 11.058 | 0.122 | 3.809 | 0.151 |

**Figure 1**

10

(a) 1-threshold      (b) 0.67-threshold      (c) 0.40-threshold

Divergence time (Million years age)

**Figure 2**

11

**(a)**

**(b)**

**(c)**

**(d)**

**Figure 3**