REVIEW ARTICLE

Ecology and Evolution
Open Access WILEY

Discriminating woody species assemblages from National Forest Inventory data based on phylogeny in Georgia

Alexander Wellenbeck[1,2] | Lutz Fehrmann[2] | Hannes Feilhauer[3] | Sebastian Schmidtlein[4] | Bernhard Misof[1,5] | Nils Hein[5]

[1]Systematic Zoology, University of Bonn, Bonn, Germany

[2]Forest Inventory and Remote Sensing, University of Göttingen, Göttingen, Germany

[3]Remote Sensing Centre for Earth System Research (RSC4Earth), Leipzig University, Leipzig, Germany

[4]Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

[5]Leibniz Institute for the Analysis of Biodiversity Change (LIB), Museum Koenig, Bonn, Germany

Correspondence
Alexander Wellenbeck, Systematic Zoology, Meckenheimer Allee 169, Bonn 53115, Germany.
Email: alex.wellenbeck@posteo.de

Abstract

Classifications of forest vegetation types and characterization of related species assemblages are important analytical tools for mapping and diversity monitoring of forest communities. The discrimination of forest communities is often based on β-diversity, which can be quantified via numerous indices to derive compositional dissimilarity between samples. This study aims to evaluate the applicability of unsupervised classification for National Forest Inventory data from Georgia by comparing two cluster hierarchies. We calculated the mean basal area per hectare for each woody species across 1059 plot observations and quantified interspecies distances for all 87 species. Following an unspuervised cluster analysis, we compared the results derived from the species-neutral dissimilarity (Bray-Curtis) with those based on the Discriminating Avalanche dissimilarity, which incorporates interspecies phylogenetic variation. Incorporating genetic variation in the dissimilarity quantification resulted in a more nuanced discrimination of woody species assemblages and increased cluster coherence. Favorable statistics include the total number of clusters (23 vs. 20), mean distance within clusters (0.773 vs. 0.343), and within sum of squares (344.13 vs. 112.92). Clusters derived from dissimilarities that account for genetic variation showed a more robust alignment with biogeographical units, such as elevation and known habitats. We demonstrate that the applicability of unsupervised classification of species assemblages to large-scale forest inventory data strongly depends on the underlying quantification of dissimilarity. Our results indicate that by incorporating phylogenetic variation, a more precise classification aligned with biogeographic units is attained. This supports the concept that the genetic signal of species assemblages reflects biogeographical patterns and facilitates more precise analyses for mapping, monitoring, and management of forest diversity.

KEYWORDS
beta diversity, community discrimination, dissimilarity, diversity monitoring, National Forest Inventory, phylogeny, Georgia, unsupervised clustering

აბსტრაქტული
ტყის მცენარეულობის ტიპების კლასიფიკაცია და მონათესავე სახეობათა შეკრების დახასიათება მნიშვნელოვანი ანალიტიკური ინსტრუმენტებია ტყის ტიპების აღწერისა და მრავალფეროვნების მონიტორინგისთვის. ტყის ტიპების განსხვავება ხშირად ემყარება β-მრავალფეროვნებას, რომლის რაოდენობრივი დადგენა შესაძლებელია მრავალი ინდექსის მეშვეობით ნიმუშებს შორის კომპოზიციური განსხვავებულობის გამოსათვლელად. ეს კვლევა მიზნად ისახავს შეაფასოს საქართველოს ეროვნული ტყის ინვენტარიზაციის ზედამხედველობის გარეშე კლასიფიკაციის გამოყენებადობა ორი კლასტერული იერარქიის შედარების გზით. ჩვენ გამოვთვალეთ საშუალო ბაზალური ფართობი ჰექტარზე თითოეული მერქნიანი სახეობისთვის 1059 ნაკვეთზე დაკვირვებით და რაოდენობრივად დავადგინეთ სახეობათაშორისი მანძილი 87-ვე სახეობისთვის. ჩვენ შევადარეთ სახეობების ნეიტრალური განსხვავებულობიდან მიღებული შედეგები (ბრეი-კურტისი) ზავის დისკრიმინაციული განსხვავებულობის საფუძველზე, რომელიც აერთიანებს სახეობათაშორის ფილოგენეტიკურ ვარიაციებს. გენეტიკური ცვალებადობის ჩართვამ განსხვავებულობის რაოდენობრივ განსაზღვრებაში გამოიწვია მერქნიანი სახეობების შეკრების უფრო ნიუანსური განსხვავება და გაზრდილი კლასტერული თანმიმდევრულობა. ხელსაყრელი სტატისტიკა მოიცავს მტევანთა საერთო რაოდენობას (23 v. 20), საშუალო მანძილს მტევნის შიგნით (0.773 vs. 0.343) და კვადრატების ჯამის ფარგლებში (344.13 vs. 112.92). განსხვავებებიდან მიღებული კლასტერებმა, რომლებიც ითვალისწინებენ გენეტიკურ ვარიაციებს, აჩვენეს უფრო მდლავრი გასწორება ბიოგეოგრაფიულ ერთეულებთან, როგორიცაა სიმაღლე და ცნობილი ჰაბიტატები. ჩვენ ვაჩვენებთ, რომ სახეობების შეკრების უკონტროლო კლასიფიკაციის გამოყენებადობა ფართომასშტაბიანი ტყის ინვენტარიზაციის მონაცემებზე მტკიცედ არის დამოკიდებული განსხვავებულობის ფუქმდებელური რაოდენობრივი განსაზღვრაზე. ჩვენი შედეგები მიუთითებენ, რომ ფილოგენეტიკური ვარიაციით, უფრო ზუსტი კლასიფიკაციაა შესაძლებელი, რომელიც შეესაბამება ბიოგეოგრაფიულ ერთეულებს. ეს ამტიცებს კონცეფციას, რომ სახეობათა შეკრების გენეტიკური სიგნალი ასახავს ბიოგეოგრაფიულ ნიმუშებს და ხელს უწყობს ტყის მრავალფეროვნების აღწერას მონიტორინგისა და მართვის უფრო ზუსტ ანალიზს.

1 | INTRODUCTION

Forest ecosystems host the largest share of terrestrial biodiversity and cover approximately one-third of the global land surface (FAO, 2020; Gillerot et al., 2021; Heym et al., 2021; Torresani et al., 2019). In light of increasing pressure on forests due to climate change and the related global loss of biodiversity, also referred to as the "the sixth mass extinction," with up to 92% of terrestrial endemics being anticipated to be negatively impacted (Manes et al., 2021), reliable approaches to assess and monitor forest diversity are required (Barnosky et al., 2011; Cowie et al., 2022; Faith, 2013; Palombo, 2021). Monitoring should include the quantification of metrics that allow the classification of ecological entities based on their specific level, or degree of biodiversity, and ultimately according to their respective conservation value, which is required by conservationists (Brooks et al., 2015; Zampiglia et al., 2019). Appropriate delineation of forest communities and characterization of related species assemblages across

taxonomic groups are important analytical tools for sensible monitoring of species diversity and biodiversity management. On broader scales, such metrics can be provided by means of vegetation classification which aims to group spatial or temporal diversity of species within a finite set of abstract categories (de Cáceres et al., 2015). Vegetation classification has proven to provide adequate means for descriptive reporting, communication, and mapping of forest communities, and related concepts have responded to changing information needs over time. Consequently, forest-type classifications exist for a wide range of targets, that is, habitat quality (qualitative assessments for biodiversity management), development over time (i.e., stand classification according to age classes for forest management) or along biogeographic gradients, and remote sensing-based mapping of ecological communities (de Cáceres et al., 2013; de Cáceres, Martín-Alcón, et al., 2019; Fassnacht et al., 2016; Hao et al., 2021).

One approach for classifying forest communities focuses on the variation in species compositional characteristics across assemblages of different sites within a geographic area, which is commonly known as β-diversity (Legendre & de Cáceres, 2013; Magurran & McGill, 2011). β-diversity can be assessed by the change in species compositional characteristics between sites (i.e., species turnover, Jost, 2010) and a plethora of metrics exist to quantify the degree of dissimilarity between assemblages on various spatial and temporal scales (de Cáceres et al., 2013; de Cáceres, Coll, et al., 2019; Legendre & Legendre, 2012; Magurran & McGill, 2011; Ricotta, 2005). The most common dissimilarity indices are exclusively based on compositional characteristics, that is, species richness and elements of evenness (Magurran, 2005), while interspecies variability (i.e., phylogenetic, taxonomic, functional, or traits) is not considered (Chao et al., 2018; Chiu et al., 2014; de Cáceres et al., 2013; Hao, Ganeshaiah, et al., 2019; Pavoine et al., 2013). In line with the increasing recognition that genetic diversity comprises an integral part of biodiversity, for example, as stated in the definition of biodiversity by the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES, Díaz et al., 2015), literature on how to incorporate phylogenetics as aspect of diversity is growing rapidly (Chao et al., 2023). Accordingly, several authors have approached forest community classification by accounting for both compositional data and interspecies phylogenetic variability (i.e., Capelo, 2020; Hao et al., 2021; Ricotta et al., 2020; Webb et al., 2002). As phylogenetically closely related species often share beneficial traits for specific environments, discriminating assemblages based on phylogenetic distances can serve as a proxy for classifying forest communities according to functional roles, environmental diversity, and conservation value (Faith, 2013; Gilbert & Parker, 2022; Hawkins et al., 2014; Padullés Cubino et al., 2021; Pavoine, 2016; Pavoine & Ricotta, 2014; Tucker et al., 2017). Hao, Ganeshaiah, et al. (2019) demonstrated that different patterns emerged if interspecies taxonomic distances were considered for the classification of global forest communities using the Discriminating Avalanche index (Ganeshaiah & Shaankar, 2000).

On a smaller scale, National Forest Inventories (NFIs) provide systematic and periodical observations of tree species abundances based on permanent sample units on a country-wide level (Corona et al., 2011). The continuous adaptation of variables assessed during NFIs highlights an increasing emphasis on aspects of biodiversity, enabling ecologists to investigate potentials and limitations of the thus provided data (Alberdi et al., 2019; Corona et al., 2011; Didion, 2020; McRoberts et al., 2009). Incorporating phylogenetic diversity of species assemblages extends the perspective on diversity in this context and bears the potential to deepen our understanding of the complex interactions among woody species over large geographical scales.

In the present study, we compare the performance of two dissimilarity indices for the discrimination of forest woody species assemblages when applied to large-scale forest inventory data such as the dataset of the first NFI of Georgia. To this end, we applied unsupervised clustering to the obtained dissimilarity matrices based on a conventional and a dissimilarity index that incorporates interspecies phylogenetic distances, respectively. Apart from statistics for internal evaluation of the resulting classifications, our comparison considered the distribution of discriminated assemblages along biogeographic gradients. Based on the assumption that genetic variability of co-inhabiting species provides a signal that sufficiently reflects site-specific environmental determinants, we investigated whether including this variable in the quantification of dissimilarity results in an improved reflection of biogeographic gradients. To test the general applicability, we incorporated the phylogenetic variability into the classification of a large, real-world dataset and evaluated the results considering cluster cohesiveness and overall interpretability.

2 | DATA AND METHODS

We compare two dissimilarity indices for the classification of woody species assemblages when applied to NFI data of Georgia. Adhering to the methodological approach underlying the data, we focus on woody species, that is, all recorded species that meet the specified target diameter at breast height (DBH, at 1.3 m, MEPA, 2018). Consequently, we are referring to woody species even if only species is written hereafter.

2.1 | Study area

Georgia is located between the Southern Slopes of the Greater and the Northern part of the Lesser Caucasus, between 41°07′ – 43°35′ N and 40°04′ – 46°44′ E (Fischer et al., 2018). The forests of Georgia host large shares of endemic species and form part of the Caucasus biodiversity hotspot (Joppa et al., 2011; Myers, 2003). Existing forest formations range from Alpine coniferous forests dominated by *Abies nordmanniana* (Steven) Spach. and *Picea orientalis* (L.) Peterm. to open juniper woodland (dominated by *Juniperus polycarpos excelsa* subsp. *polycarpos* (K. Koch) Takht. and *J. foetidissima* Willd.), encompassing further thermophilus to xerophytic mixed oak forest (*Quercus petrea* subsp. *iberica* (Steven ex M. Bieb.) Krassiln.,
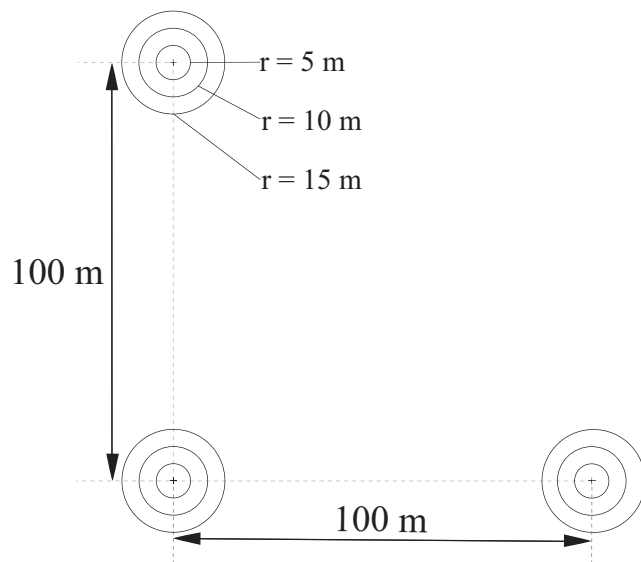
WILEY_Ecology and Evolution_
Open Access

FIGURE 1 Cluster plot configuration of the first National Forest Inventory of Georgia. Each cluster consists of three subplots for tree assessment within three concentric circles according to the measured diameter at breast height (DBH, at 1.3 m, MEPA, 2023).

Carpinus betulus L., and *C. orientalis* Mill.), Colchic alder carrs which are dominated by *Alnus glutinosa* subsp. *barbata* (C. A. Mey.) Yalt. and oriental beech (*Fagus orientalis* Lipsky) and hornbeam-oriental beech forests (Bohn et al., 2007; Dolukhanov, 2010; Fischer et al., 2018; Nakhutsrishvili et al., 2021; Novák et al., 2023).

2.2 | Forest community data

Between 2018 and 2021, Georgia implemented its' first NFI based on a systematic sampling grid of 3.6×3.6 km with a randomly selected origin. Sampling units consist of cluster plots (0.21 ha) containing three subplots of 0.0.7 ha each. These subplots are arranged in an L-shaped configuration with a distance of 100 m along both axes (Figure 1). As 18% of the country's territory is currently not accessible for government officials due to an ongoing political conflict (MEPA, 2023), clusters were sampled on approximately 74% of the national forest area (Figure 2).

2.2.1 | Assessment of Woody species

Woody species were recorded per subplot on three concentric nested subplots according to any stems' respective DBH. Stems with DBH ≥ 30 cm were recorded on the largest plot ($r = 15$ m). Stems with DBH ≥ 15 and ≥ 8 cm were recorded on the inner nested plot radii of 10 m radius and 5 m, respectively (Figure 1). Numerous variables were recorded per woody species, along with the polar coordinates of the stem axis, species, and DBH (MEPA, 2018).

From the entire NFI data set ($N = 2006$), all accessible clusters pertaining to the locally applied land use class "Forest" and

sub-classes "Tree covered area" or "Fire affected forests" were selected for analysis (MEPA, 2018). Subplots with recorded intersections with a forest boundary ("slopover sample plots") were excluded to avoid including samples with extreme outliers regarding species richness due to edge effects (Willmer et al., 2022). Clusters containing species observations that were not unambiguously identified at species level, for example, *Deciduous* spp. and *Genera* spp., were omitted because a precise quantification of the cophenetic distances along the phylogenetic hierarchy is not possible. Clusters containing subplots without woody species observations were excluded. Our sample consists of all cluster plots comprised of observations from three subplots ($n = 1059$, henceforward referred to as "samples"), which represent 53% of all clusters. Figure 2 provides an overview of the spatial distribution of samples used for the analysis.

After reprojection of sample locations to UTM38N, WGS84 (EPSG: 32638), sample elevations [m above sea level] were derived from the digital elevation model (DEM) provided by Shuttle Radar Topography Mission (SRTM, Farr et al., 2007). Elevation values were calculated as the median of all raster cell values (30×30 m) contained in or crossed by the circular subplot area ($r = 15$ m + recorded GPS error [m]) of the southwestern subplot.

2.2.2 | Species diversity data

Diversity measures are usually based on data representing the compositional variation between species (i.e., occurrence and abundance) in an assemblage (Ricotta et al., 2021). Forest species communities may display similar compositional characteristics, in terms of counts of observed species and respective individuals. However, species can be represented by large numbers of small-diameter trees, or stems belonging to the same individual, or fewer individuals but with significantly large relative shares of total basal area. Hence, abundance estimates based on counts of individuals do not take significant differences in the size structure of occurring species into consideration and may result in distinct evenness profiles. We used mean basal area (BA, m^2/ha) per species and cluster plot as species abundances to account for the variation in size of the constituents. By weighting compositional data using BA, we incorporate valuable structural information that considers site occupation per species for the quantification of β-diversity (de Cáceres, Coll, et al., 2019; McRoberts et al., 2009; Staudhammer & LeMay, 2001; Yao et al., 2019). Consequently, BA of all living stems was aggregated per cluster plot and species and divided by 3 to obtain mean BA estimates per sample (de Cáceres et al., 2015; MEPA, 2018).

2.3 | Analysis

2.3.1 | Nomenclature

Spelling and nomenclature of all recorded species were standardized with the Taxonomic Backbone databases of World Flora Online (WFO DB, Kindt, 2020) and the Global Biodiversity Information Facility (GBIF
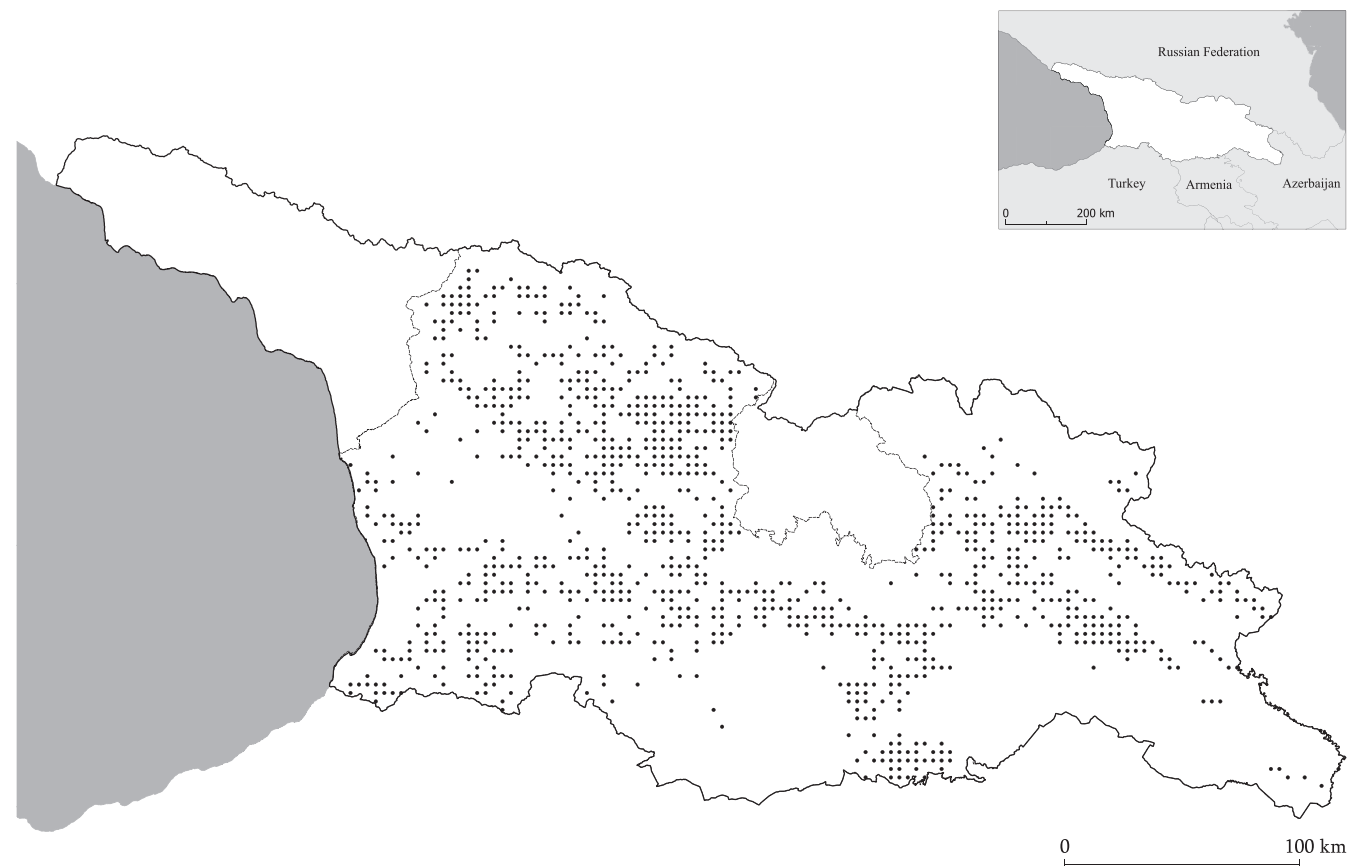
FIGURE 2 Locations of samples used from the National Forest Inventory of Georgia, $n = 1059$. Samples (black dots) consist of accessible cluster plots of equal sample size (three subplots) located inside forests that are not intersected by forest boundaries and contain only records of taxa identified at species level. Dashed lines mark the boundaries of inaccessible areas from where no field data were obtained.

Secretariat, 2021). In cases where species listed in the NFI data did not yield an unequivocal match in the WFO DB, corresponding records were harmonized with Lachashvili et al. (2022) and the nomenclature of the World Plants database (https://www.worldplants.de) to derive names for all species ranked as taxonomically accepted.

2.3.2 | Phylogenetic interspecies distances

A phylogenetic tree encompassing all observed species was constructed by matching the harmonized species list with a mega phylogeny containing 72,570 species of vascular plants according to the World Plants database (GBOTB.extended.WP.tre, Jin & Qian, 2022). The backbone mega phylogeny is based on the species-level phylogeny for vascular plants derived from gene sequencing from 7 gene regions and 39 fossil calibrations created by Zanne et al. (2014), which was updated and expanded by Jin and Qian (2022). Following the authors' recommendation to consolidate taxa below species level (e.g., sub-species) with the respective parental species, five infraspecific taxa were combined with their parental species, resulting in the lowest taxonomic unit being species level (Figure 3). From the thus created ultrametric phylogenetic tree (Jin & Qian, 2019; Qian & Jin, 2016; Smith & Brown, 2018), cophenetic distances, that is,

the total branch length connecting each pair of species at the terminal nodes of the respective phylogeny, were calculated (Bevilacqua et al., 2021; Hao, Corral-Rivas, et al., 2019; Kling et al., 2018).

Interspecies phylogenetic distances were normalized within a square matrix that contained pairwise distance values $[0 < \text{dist}_{ph} \leq 1]$ between each pair of species.

2.3.3 | Dissimilarity indices

The Discriminating Avalanche index (dA equation [2] in Table 1) developed by Ganeshaiah and Shaankar (2000) and described by Hao, Corral-Rivas, et al. (2019) considers interspecies dissimilarity by multiplying absolute differences in frequencies (numerator in BC) of species i and j in two samples with a specific distance between species i and j. We use the phylogenetic distances to weigh the difference in mean basal area between i and j. Table 1 shows the formulas of both indices used in this study.

As the maximum dissimilarity value obtained by $dA = \left(1 - \frac{1}{n}\right)$, where n equals the number of species, the resulting dissimilarities were normalized via $x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}$, with x_{min} and x_{max} representing the minimum and maximum value of dA, respectively (Hao, Corral-Rivas, et al., 2019; Legendre & Legendre, 2012). Consequently,

FIGURE 3 Phylogeny for 87 species listed in the data sample of the National Forest Inventory of Georgia. Interspecies phylogenetic distances were calculated as total branch length connecting each pair of species at the terminal nodes of the hierarchy. For respective branch lengths, see Figure S1.

pairwise dissimilarity values of the two resulting dissimilarity matrices (1059×1059) range between 0 and 1. Pairwise values of 1 imply that two samples are completely different as they do not share any species, whereas values of 0 indicate two samples are equal in terms of compositional characteristics (Chao et al., 2005; Legendre & Legendre, 2012; Leyer & Wesche, 2008). Prior to clustering, a Mantel test (Mantel, 1967) was performed to check for existing correlations between the two dissimilarity matrices. As the pairwise dissimilarities are not normally distributed, and non-linear relationships between the pairwise dissimilarity values exist, we used the

Spearman correlation coefficient with 9999 permutations (Legendre & Legendre, 2012).

2.3.4 | Isometric partitioning

The ISOPAM algorithm (Schmidtlein et al., 2010) available in package "isopam" (v. 2.0, Schmidtlein et al., 2024) combines isometric feature mapping and partitioning around medoids (data points that are most centrally located within each cluster with the sum of dissimilarities

between medoids and all other data points being minimized) in order to build clusters with a maximum number and fidelity of indicative species. The isomap ordination, which is based on geodesic distances strongly determined by neighborhood definitions, is repeated with different parameter settings. The result is clustered, and the clusters are evaluated according to the criteria mentioned above. These criteria are similar to those used when structuring phytosociological tables (Abe, 2021). In this, and the use of an ordination, Isopam is similar to Twinspan (Hill, 1979), but does not involve internal readjustments, uses geodesic distances (taking account of "neighbors of neighbors" in feature space), and works on multidimensional ordination spaces. It has been previously used for large-scale classifications of forests (Cabido et al., 2018; Černý et al., 2015; Zeballos et al., 2020) and other systems (Feilhauer et al., 2021; Hein et al., 2014; Peterka et al., 2017; Yu et al., 2022). Isopam can be run both unsupervised and supervised (with reference plots). For the current study, the original source code was extended to support dA (Capelo, 2020) and executed on a computer with two Intel Xeon CPUs (E5-2630 v3) and 256 GB RAM using R Statistical Software (v 4.3.2; R Core Team, 2023). To ease comparability, we set the maximum number of hierarchy levels to four for both classifications. Subsequently, we extracted lists of indicator species frequencies with levels of significance according to Fisher's exact test for each cluster using the Isotab function, which is part of the "isopam" package. Fidelity ("equalized phi," Tichy & Chitrý, 2006) together with Fisher's exact test if the observed frequency is not attained by chance are the criteria for qualifying as an indicator species in Isotab (Schmidtlein et al., 2024).

2.3.5 | Evaluation of clustering

To evaluate the correspondence between the original sample dissimilarities and dendrogram distances, we calculated cophenetic correlation coefficients for each hierarchical cluster structure (Lapointe & Legendre, 1995; Legendre & Legendre, 2012). The modified Rand index was used to evaluate clustering performance based on the consistency between partitions (Legendre & Legendre, 2012). Cluster homogeneity was evaluated via within sum of squares (WSS, Hao, Ganeshaiah, et al., 2019) and a comparison of the average distance between and within clusters using the function CLUSTER.STATS of the R package "fpc." To assess relevance of the hierarchies, we compared indicator species and the resulting distributions of relative BA among partitioned groups (de Cáceres et al., 2015). Evaluation of correspondence to biogeographic units was based on a comparison of elevational ranges derived from the DEM between groups and the spatial distribution of clusters in relation to forest vegetation-type classifications presented by Bohn et al. (2007). We applied the nonparametric Kruskal–Wallis test to check for significance between groups due to nonnormality in the distribution of elevation values within groups (Shapiro test). Henceforward, we are referring to the initial partition at the lowest level of the hierarchy as classes, to the intermediary partitions as branches, and the resulting clusters as assemblages.

2.3.6 | Data analysis

The analysis was conducted in R Studio version 2023.09.0-463 (RStudio Team, 2020) using R Base version 4.3.0 (R Core Team, 2023). Harmonization of nomenclature was realized via the R package "Worldflora" version 1.13-2 (Kindt, 2020), and the package "V.phylomaker2" version 0.1.0 was used to match observed species with the phylogenetic backbone (Jin & Qian, 2022). The Mantel test and BC dissimilarities were calculated using the packages "vegan" version 2.6-4 (Oksanen et al., 2022). A custom function was embedded in the adjusted clustering algorithm of the corresponding R package "isopam" version 1.2.0 for dA (Schmidtlein et al., 2022). Clustering metrics were obtained using the R packages "mclust," version 6.0.0 (Scrucca et al., 2016), and "fpc," version 2.2-10 (Hennig, 2023).

3 | RESULTS

Compositional data of $n=1059$ samples containing 65,818 living tree observations were analyzed (Table 2). In total, 87 species were represented by 52 genera, 29 families, 16 orders, and two classes.

TABLE 1 Dissimilarity indices used in this study.

Bray–Curtis dissimilarity index (1957)	$BC = \frac{\sum_{i=1}^n |p_i^a - p_i^b|}{\sum_{i=1}^n (p_i^a + p_i^b)}$	[1]
Discriminating Avalanche (Hao, Corral-Rivas, et al., 2019)	$dA = \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \Delta_i^{a,b} d_{ij} \Delta_j^{a,b}$	[2]

Note: With d_{ij} = phylogenetic distance between species i and j ($d_{ij}=d_{ji}$ and $d_{ii}=0$); $\Delta_i^{a,b}$ = absolute difference between the frequencies of species i in plots a and b ($|p_i^a - p_i^b|$); n = total number of sample plots; p_i^a, p_i^b = relative frequencies of species i in plots a and b.

TABLE 2 Summary statistics of compositional data of cluster plot observations of the National Forest Inventory of Georgia.

	Species richness				Mean basal area per species [m²/ha]			
n	Min	Max	Mean[a]	CV%	Min	Max	Mean[a]	CV%
1059	1	12	4.96 (±2.14)	43.08	2.54	79.02	30.4 (±12.79)	42.06

Abbreviations: CV%, Coefficient of variation; Var., Variance.
[a]Means are denoted with standard deviation in parenthesis.

Angiosperms were represented by 76 species (in 87.4% of all samples) across 26 different families. _Fagaceae_ (six species) accounted for the highest number of observations followed by _Betulaceae_ with 10 species (25.3% and 23.2% of all samples, respectively), _Rosaceae_, and _Sapindaceae_ (15 and 6 species, in 11.4% and 10.3% of all samples, respectively). In contrast, gymnosperms (10.3%) were represented by 11 species, belonging to three different families, with the largest family being _Pinaceae_ (6 species) followed by _Cupressaceae_ and _Taxaceae_ (4 and 1 species, respectively). Gymnosperms were observed in 33% of all samples.

3.1 | Pairwise dissimilarity

Mantel statistics ($r = .613$ with $p = .0004$) indicated a significant positive correlation between the two dissimilarity matrices. Pairwise dissimilarities according to dA ranged from 0.23 to 0.87, which is almost double the range of that of BC (0.63 to 0.99, Table 3).

Consequently, the mean dissimilarity between samples (0.79 for BC and 0.38 for dA, respectively) and thus overall variation was higher for BC than for dA, as for the latter, frequencies of dissimilarity values <1 were more evenly distributed, with very few pairwise dissimilarities of 1. Frequency distributions of sample dissimilarities are provided in Figure S2.

3.2 | Discrimination of assemblages

The hierarchical clustering based on BC (HC_{BC}) distinguished 23 assemblages (clusters) across four hierarchical levels (I–IV) and classes. Within HC_{BC}, samples were partitioned into 10 and 17 branches at levels II to III, respectively. The hierarchical clustering based on dA (HC_{dA}) contained 10 and 15 branches at levels II to III and resulted in 20 distinct assemblages over four levels and classes. Two assemblages were not partitioned further below level II in HC_{dA} (Figure 4). The number of samples per assemblage ranged from 6 to 152 for HC_{BC} (Mean: 82 ± 91.5) and 3 to 163 for HC_{dA} (Mean: 106.1 ± 119.9), respectively. For HC_{BC}, the number of observed species per assemblage ranged from 12 to 54. For HC_{dA}, one assemblage contained only four species, whereas two assemblages encompassed 53 species.

The cophenetic correlation coefficients were 0.492 and 0.442 for HC_{BC} and HC_{dA}, respectively. The obtained adjusted Rand index of 0.317 suggests a modest level of similarity between the clustering results. Average distances within clusters ranged from 0.671 to 0.875 and 0.196 to 0.49 for HC_{BC} and HC_{dA}, respectively. Mean

distance between clusters was 0.789 for HC_{BC} and 0.384 for HC_{dA}. WSS values of HC_{BC} amounted to 344.13 and 112.92 for HC_{dA}.

3.2.1 | Evaluation of cluster hierarchies

To evaluate the performance of BC and dA for clustering, we compared the resulting classes, groups, and assemblages considering internal metrics of the partitioning process. A total of 77 and 68 indicator species were listed for all partitions based on BC and dA, respectively. For a characterization based on indicators and respective frequencies, only highly significant ($p \leq .001$) species with total frequencies $\geq 50\%$ were considered, unless indicated otherwise. Indicator species for both partitions were _A. nordmanniana_, _A. cappadocicum_, _A. glutinosa_, _C. betulus_, _Carpinus orientalis_ Mill., _C. sativa_, _F. orientalis_, _P. orientalis_, _Q. petraea_ subsp. _polycarpa_ (Schur) Raus, and _Tilia rubra_ subsp. _caucasica_ (Rupr.) V. Engl. In addition, _Fraxinus excelsior_ L. is an indicator for HC_{dA}. The total number of indicators with frequencies of 100% was 24 and 8 for HC_{BC} and HC_{dA}, respectively. Class 1 of HC_{BC} is characterized by a high frequency of _F. orientalis_ (99%), whereas for HC_{dA}, _F. orientalis_ and _C. betulus_ are listed with frequencies of 93% and 75%, respectively. In class 2, the highest frequencies are observed for _Q. petrea_ (87%) and _C. betulus_ (82%). _P. orientalis_ (68%), _F. orientalis_ (57%), and _A. nordmanniana_ (53%) are the most frequent indicator species in class 3, whereas for class 4, these are _A. glutinosa_ (87%) and _Castanea sativa_ Mill. (61%). In HC_{dA}, highly significant indicators in class 1 are _F. orientalis_ (93%) and _C. betulus_ (75%), whereas in class 2, these are _Q. petrea_ (84%), _C. orientalis_ (61%), and _Fraxinus excelsior_ L. (50%). In classes 3 and 4 of HC_{dA}, _P. orientalis_ (79%), _A. nordmanniana_ (64%), _A. glutinosa_ (92%), and _C. sativa_ (59%) represent highly significant indicators in classes 3 and 4, respectively. Based on these characteristics, we labeled the four main classes according to predominant relative BA and are referring to these for ease of readability henceforward as follows: class 1 is characterized by a dominance of _Fagus_, class 2 is _Carpinus-Quercus_ dominated, and classes 3 and 4 are _Pinaceae_ and _Alnus-Castanea_ dominated, respectively. Synoptic tables of both cluster hierarchies (Figures S3 and S4, respectively) and a detailed description of indicator distributions per partition below level I for HC_{dA} (Appendix S1) are provided as appendices.

3.2.2 | Elevation and spatial distribution

To evaluate the correspondence of assemblages to existing biogeographic units, we compared the distribution of sample elevations
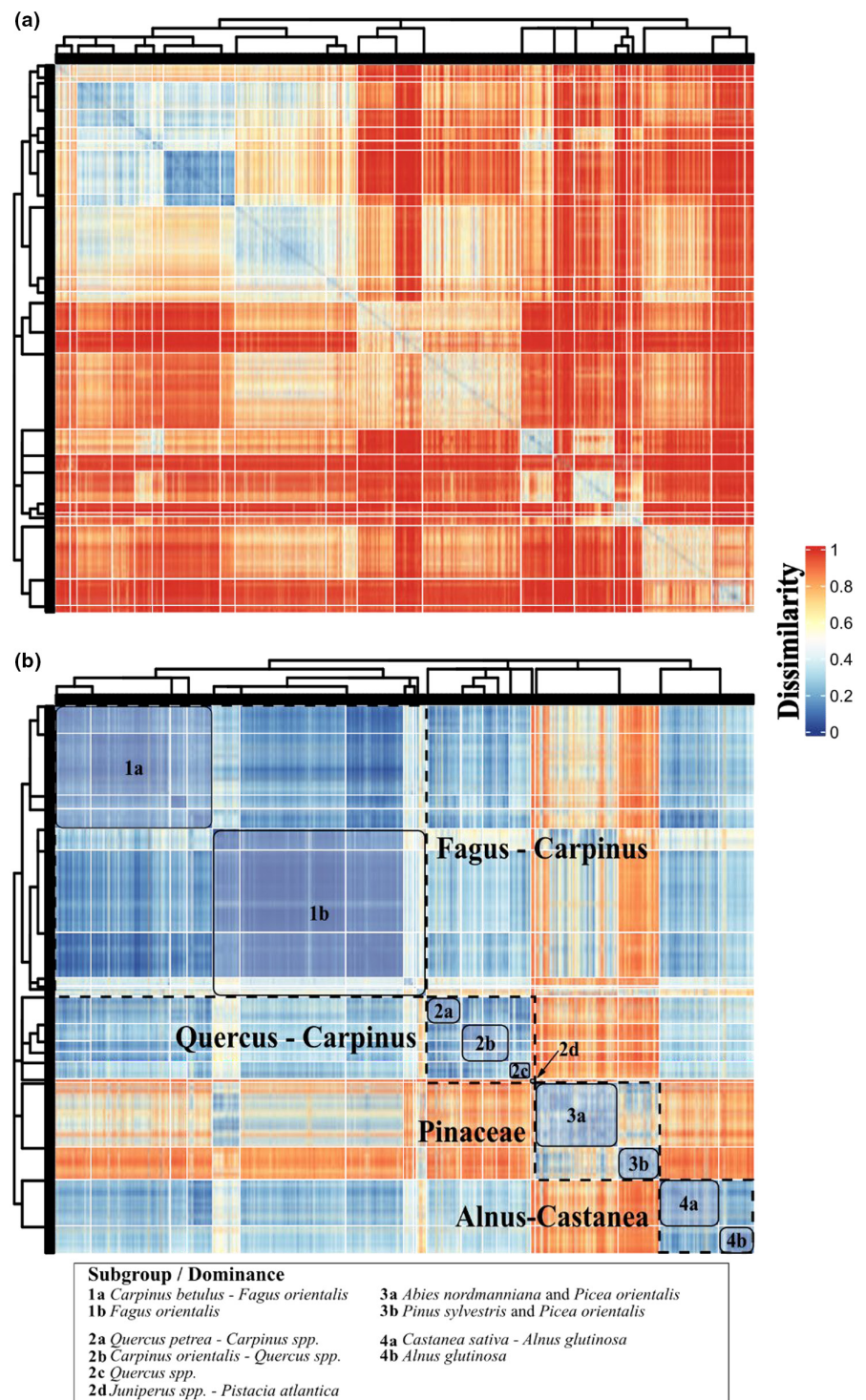
Index	n	Min	Max	Mean[a]	Var.	CV%
BC	1059	0.627	0.998	0.789 (±0.099)	0.01	12.55
dA		0.225	0.87	0.383 (±0.147)	0.022	38.38

TABLE 3 Summary statistics of mean pairwise dissimilarities between samples based on Bray–Curtis (BC) and Discriminating Avalanche (dA) of the National Forest Inventory data of Georgia.

Abbreviations: CV%, Coefficient of variation; Var., Variance.

[a]Means are denoted with standard deviation in parenthesis.

FIGURE 4 Resulting cluster hierarchy from the isopam partitioning (dendrogram) and pairwise dissimilarities of 1059 samples from the National Forest Inventory data of Georgia. Dissimilarities are based on the Bray–Curtis (a) and discriminating avalanche (b) index. The cell grid is colored according to the dissimilarity values between samples (rows and columns). In total, 23 and 20 assemblages were discriminated for Bray–Curtis and discriminating avalanche, respectively. At level II of the hierarchical clustering based on the discriminating avalanche index, 10 subgroups distributed over four classes were labeled according to dominance of relative basal area.

Subgroup / Dominance
1a *Carpinus betulus - Fagus orientalis*
1b *Fagus orientalis*

2a *Quercus petrea - Carpinus spp.*
2b *Carpinus orientalis - Quercus spp.*
2c *Quercus spp.*
2d *Juniperus spp. - Pistacia atlantica*

3a *Abies nordmanniana* and *Picea orientalis*
3b *Pinus sylvestris* and *Picea orientalis*

4a *Castanea sativa - Alnus glutinosa*
4b *Alnus glutinosa*

within assemblages. Within the four classes in both hierarchies, sample elevations are distributed similarly. Samples within the *Fagus*-dominated group (class 1) cover a wide elevational range (≤750 m to >2000 m asl), however, in HC_{BC}, 81% of all samples are located between >1000 and 2000 m asl, whereas in HC_{dA}, most samples (77%) are located within the lower range of >750 and 1750 m asl. Samples with *Carpinus-Quercus* dominance are predominantly located at elevations <1250 m in both hierarchies. The majority of *Pinaceae*-dominated samples lie above 1250 m asl, whereas all of

Alnus-Castanea-dominated samples are situated below 1250 m asl, with the majority (70%) positioned below 750 m asl for HC_{BC} and HC_{dA}, respectively. Overall classes, except for the *Fagus*-dominated group, agglomerations of sample elevation values are more pronounced within assemblages of HC_{dA} than of HC_{BC}. The applied Kruskal–Wallis test revealed highly significant differences ($p \le .001$) between assemblages for both hierarchies (Figure 5).

Spatial distributions of discriminated assemblages show a general alignment along biogeographical units. *Alnus*
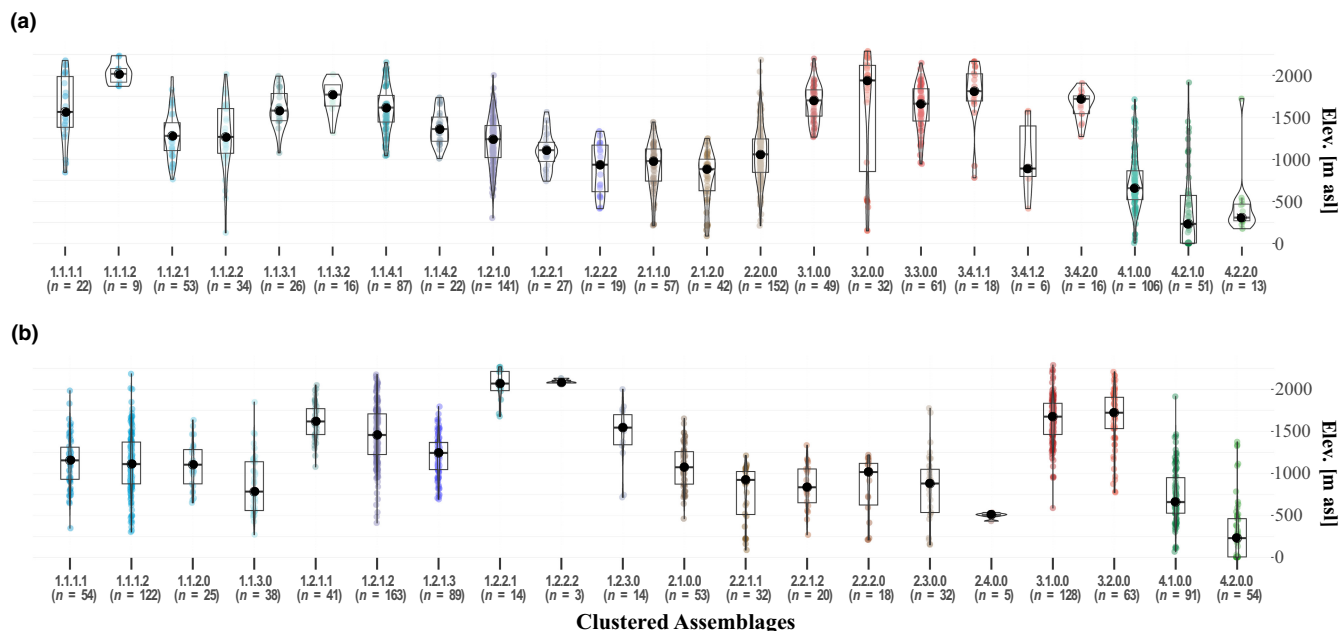
(a)

(b)

Clustered Assemblages

FIGURE 5   Sample elevations [m asl] from the National Forest Inventory in Georgia per resulting cluster for Bray–Curtis (a) and discriminating avalanche (b). The boxes represent interquartile ranges and respective median values (solid line) of sample elevations within each clustered assemblage (colored dots). Resulting p values of the $X^2_{Kruskal-Wallis}$ test are $p=2.899e-113$ and $p=1.674e-115$, with confidence intervals between 0.544, 1 and 0.538, and 1 for assemblages in the hierarchical clustering based on BC and dA, respectively ($n=1059$).

glutinosa-dominated assemblages agglomerate in the humid Alder carrs of Eastern Georgia (Nakhutsrishvili, 2013), whereas the *Pinaceae*-dominated assemblages are predominantly located in montane to subalpine areas of the Lesser and Greater Caucasus. Assemblages dominated by mixed *Quercus* spp., *Carpinus* spp., and *F. orientalis* are situated at intermediate ranges. Those with high shares of BA of *Quercus* spp. are limited to lower and drier areas, while *F. orientalis*-dominated assemblages are located at higher elevations. Interestingly, the five samples assigned to *Juniperus-Pistacia* woodland have been clearly discriminated within HC$_{dA}$ that are located in the semi-arid areas of the southwest of the country (Nakhutsrishvili et al., 2021).

To visually evaluate the spatial distribution of discriminated assemblages, we cut HC$_{dA}$ at level II because 70% of the partitions are not partitioned further below level II, resulting in 10 clustered assemblages (Figure 4). Based on characteristic indicator species, relative BA distributions, and occupied elevational zones, we labeled each assemblage accordingly and mapped the respective sample location in relation to areas of forest vegetation-type classifications according to Bohn et al. (2007) (Figure 6).

4 | DISCUSSION

Parting from the assumption that genetic variation of co-inhabiting species provides a signal that reflects site-specific environmental determinants, we contrast the performance of a species-neutral dissimilarity index (BC) with an index that considers interspecies genetic variation (dA) when used in unsupervised classification. Our

findings indicate that incorporating interspecies phylogenetic distances in the quantification of dissimilarities results in more coherent and ecologically meaningful classifications of large-scale forest inventory data with high β-diversity.

The Mantel statistics indicate a significant positive correlation between the dissimilarity matrices obtained for each index ($r=.613$, with $p=.0004$), implying that essential patterns of variation among samples are maintained in the respective quantifications. However, frequency distributions and visual inspection of dissimilarities (Figure 4) show overall higher dissimilarity values based on BC, reflecting its' sensitivity to species turnover. Whereas the resulting cluster hierarchies maintain a certain level of agreement (cophenetic correlation = .511), the fact that a correlation of one signifies complete similarity suggests that the dissimilarity signal resulting from dA is not redundant. Clustering based on dA performs slightly better in preserving the original dissimilarities according to the respective cophenetic correlation coefficients of .492 and .442 for HC$_{BC}$ and HC$_{dA}$, respectively. Dendrogram topologies, cluster validation metrics (i.e., WSS of 344.123 and 112.917 for HC$_{BC}$ and HC$_{dA}$, respectively), and distributions of relative BA among assemblages indicate a higher degree of compactness, separation, and yield, generally more conceivable clustering results based on dA. Overall, HC$_{dA}$ provided enhanced general interpretability and succeeded in discriminating clearly distinguished assemblages regarding compositional characteristics, that is, the *P. atlantica* and *Juniperus* woodlands of the semi-arid lowlands of Southeastern Georgia. These results support the concept that an extension of variables considered for quantification of dissimilarity leads to a refined conception for diversity classification if genetic variation is considered and are in line with the findings
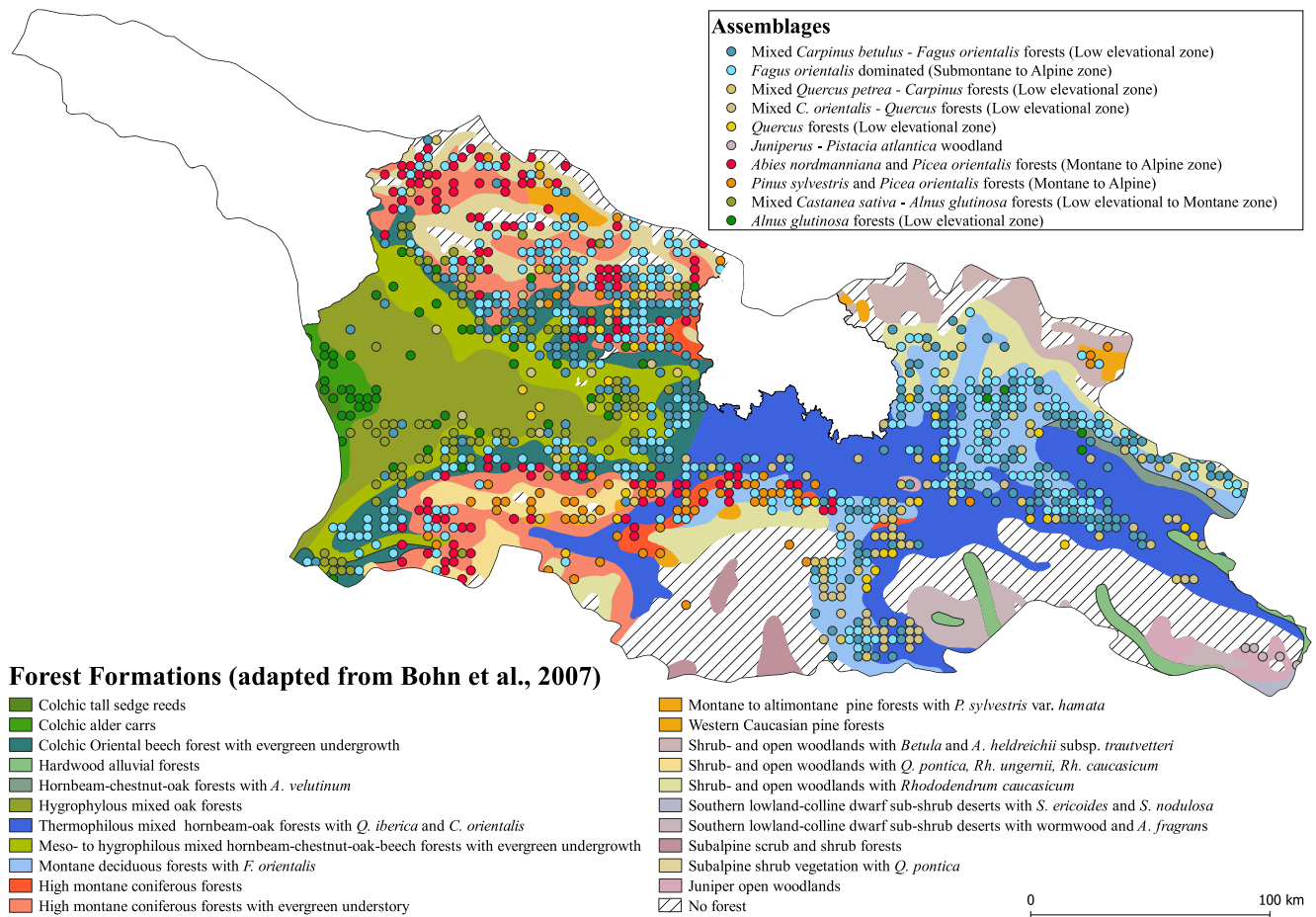
Assemblages
- Mixed *Carpinus betulus - Fagus orientalis* forests (Low elevational zone)
- *Fagus orientalis* dominated (Submontane to Alpine zone)
- Mixed *Quercus petrea - Carpinus* forests (Low elevational zone)
- Mixed *C. orientalis - Quercus* forests (Low elevational zone)
- *Quercus* forests (Low elevational zone)
- *Juniperus - Pistacia atlantica* woodland
- *Abies nordmanniana* and *Picea orientalis* forests (Montane to Alpine zone)
- *Pinus sylvestris* and *Picea orientalis* forests (Montane to Alpine)
- Mixed *Castanea sativa - Alnus glutinosa* forests (Low elevational to Montane zone)
- *Alnus glutinosa* forests (Low elevational zone)

Forest Formations (adapted from Bohn et al., 2007)
- Colchic tall sedge reeds
- Colchic alder carrs
- Colchic Oriental beech forest with evergreen undergrowth
- Hardwood alluvial forests
- Hornbeam-chestnut-oak forests with *A. velutinum*
- Hygrophylous mixed oak forests
- Thermophilous mixed hornbeam-oak forests with *Q. iberica* and *C. orientalis*
- Meso- to hygrophilous mixed hornbeam-chestnut-oak-beech forests with evergreen undergrowth
- Montane deciduous forests with *F. orientalis*
- High montane coniferous forests
- High montane coniferous forests with evergreen understory
- Montane to altimontane pine forests with *P. sylvestris* var. *hamata*
- Western Caucasian pine forests
- Shrub- and open woodlands with *Betula* and *A. heldreichii* subsp. *trautvetteri*
- Shrub- and open woodlands with *Q. pontica, Rh. ungernii, Rh. caucasicum*
- Shrub- and open woodlands with *Rhododendrum caucasicum*
- Southern lowland-colline dwarf sub-shrub deserts with *S. ericoides* and *S. nodulosa*
- Southern lowland-colline dwarf sub-shrub deserts with wormwood and *A. fragrans*
- Subalpine scrub and shrub forests
- Subalpine shrub vegetation with *Q. pontica*
- Juniper open woodlands
- No forest

0   100 km

FIGURE 6 Schematic overview of spatial distributions of discriminated assemblages and main forest formations adapted from the vegetation-type classification by Bohn et al. (2007). Assemblages are colored according to the respective palette of the four main classes of the cluster hierarchy (blue, beige, red/orange, and green for dominance of Fagus, Carpinus-Quercus, Pinaceae, and Alnus-Castanea, respectively).

of Hao, Ganeshaiah, et al. (2019) and other authors (Capelo, 2020; Pavoine & Ricotta, 2014; Ricotta & Pavoine, 2015).

As BC represents a "species-neutral" diversity index sensu Chao et al. (2010), which assumes that all observed species contribute equally to overall diversity, and species turnover constitutes the predominant signal for discrimination, reflected by the significantly higher number of indicator species with frequencies of 100% (24 and 8 for HC_{BC} and HC_{dA}, respectively). Conversely, as dA dissimilarity considers species as phylogenetic units, a complete species turnover does not necessarily result in maximum dissimilarity values between two sites because differences in abundance are weighted by the genetic proximity between species. Assuming that the genetic signal of co-occurring species reflects niche occupation within given ecogeographical areas (Hawkins et al., 2014), the thus refined dissimilarity signal appears to respond to biogeographical gradients in a more interpretable manner.

The validity of the presented approach relies on precise measurements of tree diameters and accurate species identification in the field. While traditional forest science prioritized genus-level information, growing emphasis on diversity-related issues prompted forestry experts to be increasingly trained to provide accurate species identification. The related uncertainties are not design-based issues but apply to all ecological surveys requiring botanical expert knowledge to ensure taxon detection and validate observations on species level (Lam & Kleinn, 2008; Roswell et al., 2021). Overall, only 2% of all cluster plot observations included individuals that were not identified to species level and had to be excluded. By considering only cluster plot observations of equal sample size ($m = 3$, 64% of all cluster plots) and the exclusion of subplots overlapping with the forest boundary (16%), our analyses are based on a subsample of the NFI data, representing 60% of all cluster plot observations. Hence, conclusions drawn from the presented results should consider, for example, that species exclusively occurring at forest boundaries are excluded. Potential limitations to validity arise from field sampling protocols, as overall subplot size, or sampling effects due to the nested subplot structure (with respective target DBH as inclusion criteria), may introduce bias to the quantification of dissimilarity (de Cáceres, Martín-Alcón, et al., 2019). Accordingly, observed numbers of species should be regarded as proxies of true species richness,

especially if nested sample plot designs are applied that are based on specific diameter thresholds (Lin et al., 2020). Overall species richness can be assumed to be higher with the contributions of smaller-diameter trees being neglected (Corona et al., 2017). Aggregations to cluster plot level could potentially translate to generalization effects and the loss of information on site-specific environmental factors on smaller scales. The resulting magnitude of impact on the presented results, however, is likely to vary according to forest type, topographical condition, and management regime (McRoberts et al., 2009).

With continuing advances in whole-genome phylogenetics and functional genomics, information on phylogenetic diversity is continuously improving (Kling et al., 2018). Access to comprehensive and standardized phylogenetic mega trees to quantify species genetic relationships is readily available and their application to investigate variation in community compositions is becoming increasingly more common (Gilbert & Parker, 2022; Jin & Qian, 2022).

Our results are of relevance for a wide range of classifications of ecological entities according to conservation value, mapping of ecological communities, or other discriminative objectives. The method aligns standard canopy layer data from forest inventories with natural vegetation types according to Bohn et al. (2007), but harmonization with existing forest typologies is limited due to methodological differences, such as the structural vegetation layers considered and the abundance units recorded (e.g., Chytrý et al., 2020; Mucina et al., 2016). Investigating the degree to which the resulting clusters can be aligned with defined syntaxonomical units is an interesting area for future research, especially for the identification of diagnostic species from the shrub and herb layers to refine classifications and the development of practical assessment procedures.

The integration of genetic signals of forest communities during characterization has wide implications for respective approaches to classification. As a proxy indicative of the relationship between species composition and site conditions, interspecies genetic variation extends the scope of forest diversity mapping, management, and monitoring to account for alterations inconceivable by conventional compositional variables. Beyond the potential advantage of streamlining processes by applying unsupervised classification to large datasets, our approach is straightforward and can be readily replicated with comparable data, provided entities are assessed in a systematic manner. Ecological studies are frequently less systematic and constrained by temporal and spatial scales due to the dynamic nature of communities over time and space. This holds true to a lesser degree for assessments of woody species communities within the context of national forest monitoring systems, which are resampled in fixed intervals. Hence, from a practical point of view, the resulting network of observational studies provides a valuable framework for a systematic and recurring collection of ecological data, as additional costs and efforts can be embedded into existing structures. Extending the scope of study objectives to systematic assessments of a wider range of botanical and potentially zoological taxa could provide powerful and statistically robust data for analyses of organisms, structural components, and the interrelationships between them.

5 | CONCLUSIONS

We present an approach to discriminate species diversity from NFI data of forest communities with high β-diversity and species turnover. The novelty of the method lies in considering interspecies genetic variability for the quantification of diversity and subsequent classification using an unsupervised clustering algorithm on a country-wide scale. We demonstrate that large-scale forest inventory data can be classified in an ecologically meaningful manner based on mean basal area estimates per species and consideration of phylogenetic dissimilarity between samples. The thus obtained discrimination of species assemblages provides a differentiated picture of existing diversity patterns along expected biogeographical gradients without the need for additional assessments. This approach aligns with a biodiversity concept considering genetic diversity and could potentially be standardized for application to similar datasets, provided systematic data assessment is granted. The presented results should be considered as a step in evaluating to which extent large-area forest inventory data could provide a backbone for extended biodiversity monitoring systems, as discrimination of woody species assemblages allows for systematic delineation of forest ecosystems if genetic variation is considered.

AUTHOR CONTRIBUTIONS

Alexander Wellenbeck: Conceptualization (lead); data curation (lead); formal analysis (lead); investigation (lead); methodology (lead); validation (lead); visualization (lead); writing – original draft (lead); writing – review and editing (lead). **Lutz Fehrmann:** Conceptualization (supporting); formal analysis (supporting); methodology (supporting); supervision (supporting); visualization (supporting); writing – review and editing (supporting). **Hannes Feilhauer:** Formal analysis (supporting); methodology (supporting); writing – review and editing (supporting). **Sebastian Schmidtlein:** Methodology (supporting); software (supporting); writing – review and editing (supporting). **Bernhard Misof:** Methodology (supporting); supervision (supporting); writing – review and editing (supporting). **Nils Hein:** Conceptualization (supporting); investigation (supporting); methodology (supporting); project administration (lead); writing – original draft (supporting); writing – review and editing (supporting).

CONFLICT OF INTEREST STATEMENT

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Dryad at https://doi.org/10.5061/dryad.wpzgmsbvv.

ORCID

Alexander Wellenbeck https://orcid.org/0009-0008-7007-3922

Nils Hein https://orcid.org/0000-0002-5172-8531

REFERENCES

Abe, S. (2021). Comparison of traditional and automated approaches in classification of Japanese coastal beach and dune vegetation. *Society of Vegetation Science*, 38, 67–80. https://doi.org/10.15031/vegsci.38.67

Alberdi, I., Nunes, L., Kovac, M., Bonheme, I., Cañellas, I., Rego, F. C., Dias, S., Duarte, I., Notarangelo, M., Rizzo, M., & Gasparini, P. (2019). The conservation status assessment of Natura 2000 forest habitats in Europe: capabilities, potentials and challenges of national forest inventories data. *Annals of Forest Science*, 76(2), 1–15. https://doi.org/10.1007/s13595-019-0820-4

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., Marshall, C., McGuire, J. L., Lindsey, E. L., Maguire, K. C., Mersey, B., & Ferrer, E. A. (2011). Has the Earth's sixth mass extinction already arrived? *Nature*, 471(7336), 51–57. https://doi.org/10.1038/nature09678

Bevilacqua, S., Anderson, M. J., Ugland, K. I., Somerfield, P. J., & Terlizzi, A. (2021). The use of taxonomic relationships among species in applied ecological research: Baseline, steps forward and future challenges. *Austral Ecology*, 46(6), 950–964. https://doi.org/10.1111/aec.13061

Bohn, U., Zazanashhvili, N., & Nakhutsrishvili, G. (2007). The map of the natural vegetation of Europe and its application in the Caucasus ecoregion. *Bulletin of the Georgian National Academy of Science*, 175, 112–119. http://science.org.ge/old/moambe/2007-vol1/112-120.pdf

Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland Forest communities of southern Wisconsin. *Ecological Monographs*, 27(4), 325–349. https://doi.org/10.2307/1942268

Brooks, T. M., Cuttelod, A., Faith, D. P., Garcia-Moreno, J., Langhammer, P., & Pérez-Espona, S. (2015). Why and how might genetic and phylogenetic diversity be reflected in the identification of key biodiversity areas? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 370(1662), 20140019. https://doi.org/10.1098/rstb.2014.0019

Cabido, M., Zeballos, S. R., Zak, M., Carranza, M. L., Giorgis, M. A., Cantero, J. J., & Acosta, A. T. R. (2018). Native woody vegetation in central Argentina: Classification of Chaco and Espinal forests. *Applied Vegetation Science*, 21(2), 298–311. https://doi.org/10.1111/avsc.12369

Capelo, J. (2020). Using species abundance and phylogeny conjointly to approach vegetation classification: A case study on Macaronesia's woody vegetation. *Journal of Vegetation Science*, 31(4), 616–633. https://doi.org/10.1111/jvs.12886

Černý, T., Kopecký, M., Petřík, P., Song, J.-S., Šrůtek, M., Valachovič, M., Altman, J., & Doležal, J. (2015). Classification of Korean forests: Patterns along geographic and environmental gradients. *Applied Vegetation Science*, 18(1), 5–22. https://doi.org/10.1111/avsc.12124

Chao, A., Chazdon, R. L., Colwell, R. K., & Shen, T.-J. (2005). A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters*, 8(2), 148–159. https://doi.org/10.1111/j.1461-0248.2004.00707.x

Chao, A., Chiu, C.-H., & Jost, L. (2010). Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365(1558), 3599–3609. https://doi.org/10.1098/rstb.2010.0272

Chao, A., Chiu, C.-H., Villéger, S., Sun, I.-F., Thorn, S., Lin, Y.-C., Chiang, J.-M., & Sherwin, W. B. (2018). An attribute-diversity approach to functional diversity, functional beta diversity, and related (dis)similarity measures. *Ecological Monographs*, 89(2), e01343. https://doi.org/10.1002/ecm.1343

Chao, A., Thorn, S., Chiu, C.-H., Moyes, F., Hu, K.-H., Chazdon, R. L., Wu, J., Magnago, L. F. S., Dornelas, M., Zelený, D., Colwell, R. K., & Magurran, A. E. (2023). Rarefaction and extrapolation with beta diversity under a framework of H ill numbers: The iNEXT. beta3D standardization. *Ecological Monographs*, 93, e1588. https://doi.org/10.1002/ecm.1588

Chiu, C.-H., Jost, L., & Chao, A. (2014). Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs*, 84(1), 21–44. https://doi.org/10.1890/12-0960.1

Chytrý, M., Tichý, L., Hennekens, S. M., Knollová, I., Janssen, J. A. M., Rodwell, J. S., Peterka, T., Marcenò, C., Landucci, F., Danihelka, J., Hájek, M., Dengler, J., Novák, P., Zukal, D., Jiménez-Alfaro, B., Mucina, L., Abdulhak, S., Aćić, S., Agrillo, E., … Schaminée, J. H. J. (2020). EUNIS habitat classification: Expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science*, 23(4), 648–675. https://doi.org/10.1111/avsc.12519

Corona, P., Chirici, G., McRoberts, R. E., Winter, S., & Barbati, A. (2011). Contribution of large-scale forest inventories to biodiversity assessment and monitoring. *Forest Ecology and Management*, 262(11), 2061–2069. https://doi.org/10.1016/j.foreco.2011.08.044

Corona, P., Franceschi, S., Pisani, C., Portoghesi, L., Mattioli, W., & Fattorini, L. (2017). Inference on diversity from forest inventories: A review. *Biodiversity and Conservation*, 26(13), 3037–3049. https://doi.org/10.1007/s10531-015-1017-2

Cowie, R. H., Bouchet, P., & Fontaine, B. (2022). The sixth mass extinction: Fact, fiction or speculation? *Biological Reviews of the Cambridge Philosophical Society*, 97(2), 640–663. https://doi.org/10.1111/brv.12816

de Cáceres, M., Chytrý, M., Agrillo, E., Attorre, F., Botta-Dukát, Z., Capelo, J., Czúcz, B., Dengler, J., Ewald, J., Faber-Langendoen, D., Feoli, E., Franklin, S. B., Gavilán, R., Gillet, F., Jansen, F., Jiménez-Alfaro, B., Krestov, P., Landucci, F., Lengyel, A., … Wiser, S. K. (2015). A comparative framework for broad-scale plot-based vegetation classification. *Applied Vegetation Science*, 18(4), 543–560. https://doi.org/10.1111/avsc.12179

de Cáceres, M., Coll, L., Legendre, P., Allen, R. B., Wiser, S. K., & Fortin, M.-J. (2019). Trajectory analysis in community ecology. *Ecological Monographs*, 89(2), e01350. https://doi.org/10.1002/ecm.1350

de Cáceres, M., Legendre, P., & He, F. (2013). Dissimilarity measurements and the size structure of ecological communities. *Methods in Ecology and Evolution*, 4(12), 1167–1177. https://doi.org/10.1111/2041-210X.12116

de Cáceres, M., Martín-Alcón, S., González-Olabarria, J. R., & Coll, L. (2019). A general method for the classification of forest stands using species composition and vertical and horizontal structure. *Annals of Forest Science*, 76(2), 1–19. https://doi.org/10.1007/s13595-019-0824-0

Díaz, S., Demissew, S., Carabias, J., Joly, C., Lonsdale, M., Ash, N., Larigauderie, A., Adhikari, J. R., Arico, S., Báldi, A., Bartuska, A., Baste, I. A., Bilgin, A., Brondizio, E., Chan, K. M. A., Figueroa, V.

E., Duraiappah, A., Fischer, M., Hill, R., … Zlatanova, D. (2015). The IPBES conceptual framework — Connecting nature and people. *Current Opinion in Environmental Sustainability*, 14, 1–16. https://doi.org/10.1016/j.cosust.2014.11.002

Didion, M. (2020). Extending harmonized national forest inventory herb layer vegetation cover observations to derive comprehensive biomass estimates. *Forest Ecosystems*, 7(1), 1–14. https://doi.org/10.1186/s40663-020-00230-7

Dolukhanov, A. G. (2010). *Лесная растительность Грузии: (Forest vegetation of Georgia)*. Universal Tbilisi.

Faith, D. P. (2013). Biodiversity and evolutionary history: Useful extensions of the PD phylogenetic diversity assessment framework. *Annals of the New York Academy of Sciences*, 1289, 69–89. https://doi.org/10.1111/nyas.12186

FAO. (2020). *Global Forest resources assessment 2020: Key findings*. FAO. https://doi.org/10.4060/ca8753en

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., & Alsdorf, D. (2007). The shuttle radar topography Mission. *Reviews of Geophysics*, 45(2), RG2004. https://doi.org/10.1029/2005RG000183

Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., Straub, C., & Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, 186, 64–87. https://doi.org/10.1016/j.rse.2016.08.013

Feilhauer, H., Zlinszky, A., Kania, A., Foody, G. M., Doktor, D., Lausch, A., & Schmidtlein, S. (2021). Let your maps be fuzzy!—Class probabilities and floristic gradients as alternatives to crisp mapping for remote sensing of vegetation. *Remote Sensing in Ecology and Conservation*, 7(2), 292–305. https://doi.org/10.1002/rse2.188

Fischer, E., Gröger, A., & Lobin, W. (2018). *Illustrated field guide to the flora of Georgia (South Caucasus) (1st edition). Koblenz geographical colloquia series biogeographical monographs: Vol. 3*. University of Koblenz-Landau.

Ganeshaiah, K. N., & Shaankar, R. U. (2000). Measuring biological heterogeneity of forest vegetation types: Avalanche index as an estimate of biological diversity. *Biodiversity and Conservation*, 9, 953–963. https://doi.org/10.1023/A:1008910918751

GBIF Secretariat. (2021). *Gbif backbone taxonomy*. https://doi.org/10.15468/39OMEI

Gilbert, G. S., & Parker, I. M. (2022). Phylogenetic distance metrics for studies of focal species in communities: Quantiles and cumulative curves. *Diversity*, 14(7), 521. https://doi.org/10.3390/d14070521

Gillerot, L., Grussu, G., Condor-Golec, R., Tavani, R., Dargush, P., & Attorre, F. (2021). Progress on incorporating biodiversity monitoring in REDD+ through national forest inventories. *Global Ecology and Conservation*, 32, e01901. https://doi.org/10.1016/j.gecco.2021.e01901

Hao, M., Corral-Rivas, J., González-Elizondo, M. S., Ganeshaiah, K. N., Nava-Miranda, M. G., Zhang, C., Zhao, X., & von Gadow, K. (2019). Assessing biological dissimilarities between five forest communities. *Forest Ecosystems*, 6(1), 1–8. https://doi.org/10.1186/s40663-019-0188-9

Hao, M., Gadow, K., Alavi, S. J., Álvarez-González, J. G., Baluarte-Vásquez, J. R., Corral-Rivas, J., Hui, G., Korol, M., Kumar, R., Liang, J., Meyer, P., Remadevi, O. K., Kakkar, R., Liu, W., Zhao, X., & Zhang, C. (2021). A classification of woody communities based on biological dissimilarity. *Applied Vegetation Science*, 24(1), e12565. https://doi.org/10.1111/avsc.12565

Hao, M., Ganeshaiah, K. N., Zhang, C., Zhao, X., & von Gadow, K. (2019). Discriminating among forest communities based on taxonomic, phylogenetic and trait distances. *Forest Ecology and Management*, 440, 40–47. https://doi.org/10.1016/j.foreco.2019.03.006

Hawkins, B. A., Rueda, M., Rangel, T. F., Field, R., Diniz-Filho, J. A. F., & Linder, P. (2014). Community phylogenetics at the biogeographical scale: Cold tolerance, niche conservatism and the structure of north American forests. *Journal of Biogeography*, 41(1), 23–38. https://doi.org/10.1111/jbi.12171

Hein, N., Feilhauer, H., Finch, O.-D., Schmidtlein, S., & Löffler, J. (2014). Snow cover determines the ecology and biogeography of spiders (Araneae) in alpine tundra ecosystems. *Erdkunde*, 68, 157–172. https://doi.org/10.3112/erdkunde.2014.03.01

Hennig, C. (2023). *Fpc: Flexible procedures for clustering* (version 2.2-10,) [computer software]. https://CRAN.R-project.org/package=fpc

Heym, M., Uhl, E., Moshammer, R., Dieler, J., Stimm, K., & Pretzsch, H. (2021). Utilising forest inventory data for biodiversity assessment. *Ecological Indicators*, 121, 107196. https://doi.org/10.1016/j.ecolind.2020.107196

Hill, M. O. (1979). *TWINSPAN – A FORTRAN program for arranging multivariate data in an ordered two way table by classification of individuals and attributes* (p. 48). Cornell University.

Jin, Y., & Qian, H. (2019). V.PhyloMaker: An R package that can generate very large phylogenies for vascular plants. *Ecography*, 42(8), 1353–1359. https://doi.org/10.1111/ecog.04434

Jin, Y., & Qian, H. (2022). V.Phylomaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. *Plant Diversity*, 44(4), 335–339. https://doi.org/10.1016/j.pld.2022.05.005

Joppa, L. N., Roberts, D. L., Myers, N., & Pimm, S. L. (2011). Biodiversity hotspots house most undiscovered plant species. *Proceedings of the National Academy of Sciences of the United States of America*, 108(32), 13171–13176. https://doi.org/10.1073/pnas.1109389108

Jost, L. (2010). The relation between evenness and diversity. *Diversity*, 2(2), 207–232. https://doi.org/10.3390/d2020207

Kindt, R. (2020). *WorldFlora: An R package for exact and fuzzy matching of plant names against the World Flora Online Taxonomic Backbone data*. https://doi.org/10.1101/2020.02.02.930719

Kling, M. M., Mishler, B. D., Thornhill, A. H., Baldwin, B. G., & Ackerly, D. D. (2018). Facets of phylodiversity: Evolutionary diversification, divergence and survival as conservation targets. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 374(1763), 20170397. https://doi.org/10.1098/rstb.2017.0397

Lachashvili, N., Kereselidze, K., & Kikvidze, M. (2022). The checklist of trees of Georgia (Caucasus) and their provisional regional assessment according to IUCN categories and criteria. *Flora Mediterranea*, 32, 149–188. https://doi.org/10.7320/FlMedit32.149

Lam, T. Y., & Kleinn, C. (2008). Estimation of tree species richness from large area forest inventory data: Evaluation and comparison of jackknife estimators. *Forest Ecology and Management*, 255(3–4), 1002–1010. https://doi.org/10.1016/j.foreco.2007.10.007

Lapointe, F.-J., & Legendre, P. (1995). Comparison tests for dendrograms: A comparative evaluation. *Journal of Classification*, 12, 265–282. https://doi.org/10.1007/BF03040858

Legendre, P., & de Cáceres, M. (2013). Beta diversity as the variance of community data: Dissimilarity coefficients and partitioning. *Ecology Letters*, 16(8), 951–963. https://doi.org/10.1111/ele.12141

Legendre, P., & Legendre, L. (2012). *Numerical ecology. Third english edition (third). Developments in environmental modelling: Vol. 24*. Elsevier BV.

Leyer, I., & Wesche, K. (2008). *Multivariate Statistik in der Ökologie*. Springer Berlin Heidelberg.

Lin, H.-T., Lam, T. Y., von Gadow, K., & Kershaw, J. A. (2020). Effects of nested plot designs on assessing stand attributes, species diversity, and spatial forest structures. *Forest Ecology and Management*, 457, 117658. https://doi.org/10.1016/j.foreco.2019.117658

Magurran, A. E. (2005). Biological diversity. *Current Biology*, 15(4), 116–118. https://doi.org/10.1016/j.cub.2005.02.006

Magurran, A. E., & McGill, B. J. (2011). *Biological diversity: Frontiers in measurement and assessment*. Oxford University Press.

Manes, S., Costello, M. J., Beckett, H., Debnath, A., Devenish-Nelson, E., Grey, K.-A., Jenkins, R., Khan, T. M., Kiessling, W., Krause, C.,

Maharaj, S. S., Midgley, G. F., Price, J., Talukdar, G., & Vale, M. M. (2021). Endemism increases species' climate change risk in areas of global biodiversity importance. *Biological Conservation*, *257*, 109070. https://doi.org/10.1016/j.biocon.2021.109070

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*(2), 209–220.

McRoberts, R. E., Tomppo, E., Schadauer, K., Vidal, C., Ståhl, G., Chirici, G., Lanz, A., Cienciala, E., Winter, S., & Smith, B. (2009). Harmonizing National Forest Inventories. *Journal of Forestry*, *107*(4), 179–187. https://doi.org/10.1093/jof/107.4.179

MEPA. (2018). *Field manual for the Georgian National Forest Inventory: Part II*. Ministry of Environmental Protection and Agriculture.

MEPA. (2023). ანგარიში ტყის პირველი ეროვნული აღრიცხვა საქართველოში – *First National Forest Inventory in Georgia: Report 2023*. Ministry of Environmental Protection and Agriculture. https://mepa.gov.ge/En/Files/Download/53934

Mucina, L., Bültmann, H., Dierßen, K., Theurillat, J.-P., Raus, T., Čarni, A., Šumberová, K., Willner, W., Dengler, J., García, R. G., Chytrý, M., Hájek, M., Di Pietro, R., Iakushenko, D., Pallas, J., Daniëls, F. J., Bergmeier, E., Santos Guerra, A., Ermakov, N., ... Tichý, L. (2016). Vegetation of Europe: hierarchical floristic classification system of vascular plant, bryophyte, lichen, and algal communities. *Applied Vegetation Science*, *19*(S1), 3–264. https://doi.org/10.1111/avsc.12257

Myers, N. (2003). Biodiversity hotspots revisited. *Bioscience*, *53*(10), 916–917. https://doi.org/10.1641/0006-3568(2003)053[0916:BHR]2.0.CO;2

Nakhutsrishvili, G. (Ed.). (2013). *The vegetation of Georgia (South Caucasus)*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-29915-5

Nakhutsrishvili, G., Abdaladze, O., & Batsatsashvili, K. (2021). Ecological gradients (west-east) and vegetation of the central great Caucasus. *Bocconea*, *29*, 157–168. https://doi.org/10.7320/Bocc29.157

Novák, P., Kalníková, V., Szokala, D., Aleksanyan, A., Batsatsashvili, K., Fayvush, G., Kolbaia, S., Nakhutsrishvili, G., Sedláček, V., Štěrba, T., & Zukal, D. (2023). Transcaucasian vegetation database – A phytosociological database of the southern Caucasus. *Vegetation Classification and Survey*, *4*, 231–240. https://doi.org/10.3897/VCS.105521

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Solymos, P., & Stevens, M. (2022). *vegan* (Version 2.6–4) [Computer software]. https://CRAN.R-project.org/package=vegan

Padullés Cubino, J., Lososová, Z., Bonari, G., Agrillo, E., Attorre, F., Bergmeier, E., Biurrun, I., Campos, J. A., Čarni, A., Ćuk, M., de Sanctis, M., Indreica, A., Jiménez-Alfaro, B., Khanina, L., Knollová, I., Lenoir, J., Pielech, R., Rašomavičius, V., Škvorc, Ž., ... Chytrý, M. (2021). Phylogenetic structure of European forest vegetation. *Journal of Biogeography*, *48*(4), 903–916. https://doi.org/10.1111/jbi.14046

Palombo, M. R. (2021). Thinking about the biodiversity loss in this changing world. *Geosciences*, *11*(9), 370. https://doi.org/10.3390/geosciences11090370

Pavoine, S. (2016). A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, *125*(12), 1719–1732. https://doi.org/10.1111/oik.03262

Pavoine, S., Gasc, A., Bonsall, M. B., & Mason, N. W. (2013). Correlations between phylogenetic and functional diversity: Mathematical artefacts or true ecological and evolutionary processes? *Journal of Vegetation Science*, *24*(5), 781–793. https://doi.org/10.1111/jvs.12051

Pavoine, S., & Ricotta, C. (2014). Functional and phylogenetic similarity among communities. *Methods in Ecology and Evolution*, *5*(7), 666–675. https://doi.org/10.1111/2041-210X.12193

Peterka, T., Hájek, M., Jiroušek, M., Jiménez-Alfaro, B., Aunina, L., Bergamini, A., Dítě, D., Felbaba-Klushyna, L., Graf, U., Hájková, P., Hettenbergerová, E., Ivchenko, T. G., Jansen, F., Koroleva, N. E., Lapshina, E. D., Lazarević, P. M., Moen, A., Napreenko, M. G., Pawlikowski, P., ... Chytrý, M. (2017). Formalized classification of European fen vegetation at the alliance level. *Applied Vegetation Science*, *20*(1), 124–142. https://doi.org/10.1111/avsc.12271

Qian, H., & Jin, Y. (2016). An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *Journal of Plant Ecology*, *9*(2), 233–239. https://doi.org/10.1093/jpe/rtv047

R Core Team. (2023). *R: A language and environment for statistical computing [computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/

Ricotta, C. (2005). Through the jungle of biological diversity. *Acta Biotheoretica*, *53*, 29–38. https://doi.org/10.1007/s10441-005-7001-6

Ricotta, C., Laroche, F., Szeidl, L., & Pavoine, S. (2020). From alpha to beta functional and phylogenetic redundancy. *Methods in Ecology and Evolution*, *11*(4), 487–493. https://doi.org/10.1111/2041-210X.13353

Ricotta, C., & Pavoine, S. (2015). Measuring similarity among plots including similarity among species: An extension of traditional approaches. *Journal of Vegetation Science*, *26*(6), 1061–1067. https://doi.org/10.1111/jvs.12329

Ricotta, C., Szeidl, L., & Pavoine, S. (2021). Towards a unifying framework for diversity and dissimilarity coefficients. *Ecological Indicators*, *129*, 107971. https://doi.org/10.1016/j.ecolind.2021.107971

Roswell, M., Dushoff, J., & Winfree, R. (2021). A conceptual guide to measuring species diversity. *Oikos*, *130*(3), 321–338. https://doi.org/10.1111/oik.07202

RStudio Team. (2020). *RStudio: Integrated development for R (Version 2023.09.0) [Computer software]*. Posit Software, P. B.C.

Schmidtlein, S., Collison, J., & Pfannendoerfer, R. (2024). *Isopam: Clustering of sites with species data* [CRAN]. https://CRAN.R-project.org/package=isopam

Schmidtlein, S., Jason, C., Pfannendoerfer, R., & Tichý, L. (2022). *Package 'isopam' (Version 1.1.9) [Computer software]*. Schmidtlein, Sebastian.

Schmidtlein, S., Tichý, L., Feilhauer, H., & Faude, U. (2010). A brute-force approach to vegetation classification. *Journal of Vegetation Science*, *21*, 1162–1171. https://doi.org/10.1088/0004-6256/141/2/44

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). Mclust 5: Clustering, classification and density estimation using gaussian finite mixture models. *The R Journal*, *8*(1), 289–317. https://doi.org/10.32614/RJ-2016-021

Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany*, *105*(3), 302–314. https://doi.org/10.1002/ajb2.1019

Staudhammer, C. L., & LeMay, V. M. (2001). Introduction and evaluation of possible indices of stand structural diversity. *Canadian Journal of Forest Research*, *31*(7), 1105–1115. https://doi.org/10.1139/cjfr-31-7-1105

Tichy, L., & Chytry, M. (2006). Statistical determination of diagnostic species for site groups of unequal size. *Journal of Vegetation Science*, *17*(6), 809–818. https://doi.org/10.1111/j.1654-1103.2006.tb02504.x

Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Marcantonio, M., Ricotta, C., & Tonon, G. (2019). Estimating tree species diversity from space in an alpine conifer forest: The Rao's Q diversity index meets the spectral variation hypothesis. *Ecological Informatics*, *52*, 26–34. https://doi.org/10.1016/j.ecoinf.2019.04.001

Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., Grenyer, R., Helmus, M. R., Jin, L. S., Mooers, A. O., Pavoine, S., Purschke, O., Redding, D. W., Rosauer, D. F., Winter, M., & Mazel, F. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews of the Cambridge Philosophical Society*, *92*(2), 698–715. https://doi.org/10.1111/brv.12252

Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, *33*(1), 475–505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448

Willmer, J. N. G., Püttker, T., & Prevedello, J. A. (2022). Global impacts of edge effects on species richness. *Biological Conservation*, *272*, 109654. https://doi.org/10.1016/j.biocon.2022.109654

Yao, J., Zhang, C., de Cáceres, M., Legendre, P., & Zhao, X. (2019). Variation in compositional and structural components of community assemblage and its determinants. *Journal of Vegetation Science*, *30*(2), 257–268. https://doi.org/10.1111/jvs.12708

Yu, H., Fang, G., Rose, K. A., Tang, Y., & Song, X. (2022). Examining epibenthic assemblages associated with artificial reefs using a species archetype approach. *Marine and Coastal Fisheries*, *14*(3), e10206. https://doi.org/10.1002/mcf2.10206

Zampiglia, M., Bisconti, R., Maiorano, L., Aloise, G., Siclari, A., Pellegrino, F., Martino, G., Pezzarossa, A., Chiocchio, A., Martino, C., Nascetti, G., & Canestrelli, D. (2019). Drilling down hotspots of intraspecific diversity to bring them into on-ground conservation of threatened species. *Frontiers in Ecology and Evolution*, *7*, 205. https://doi.org/10.3389/fevo.2019.00205

Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., McGlinn, D. J., O'Meara, B. C., Moles, A. T., Reich, P. B., Royer, D. L., Soltis, D. E., Stevens, P. F., Westoby, M., Wright, I. J., Aarssen, L., Bertin, R. I., Calaminus, A., Govaerts, R., … Beaulieu, J. M. (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature*, *506*(7486), 89–92. https://doi.org/10.1038/nature12872

Zeballos, S. R., Giorgis, M. A., Cabido, M. R., Acosta, A. T., Del Iglesias, M. R., & Cantero, J. J. (2020). The lowland seasonally dry subtropical forests in central Argentina: Vegetation types and a call for conservation. *Vegetation Classification and Survey*, *1*, 87–102. https://doi.org/10.3897/VCS/2020/38013

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.