Referee #2

The manuscript introduces the HDM-Plot dataset, compiled from four field campaigns conducted in 2022–2024 in the Hengduan Mountains of southwestern China, which is a global biodiversity hotspot. The geographic and elevational coverage is impressive, and the data structure appears rich enough to support vegetation classification, conservation planning, and ecological analyses. The inclusion of raw plot data, species lists, importance values, and classification tables is a strong asset. I acknowledge the difficulty to conduct field work in this region and appreciate the authors' effort in making such dataset publicly available. However, I still have a few concerns and think that the preprint and the usability of the dataset would be improved by the following comments.

Response: We sincerely thank for your review on our manuscript. According to your positive and detailed comments, we have made some changes to our manuscript. Please refer to replies as following.

Comment 1 Plot-size heterogeneity needs more clarification. The authors note that "plot size was determined following community physiognomy and stand heterogeneity," with forest plots typically 10 m × 10 m to 20 m × 20 m, shrubland plots 2 m × 2 m to 10 m × 10 m, and grassland plots 1 m × 1 m to 2 m × 2 m. That is a very large difference in sampling area among vegetation types. It would therefore be helpful to explain more explicitly how this was considered in the analyses. Please clarify whether any standardization or rarefaction was used for species richness comparisons, and whether plot-size distributions in space, along elevation, or for different vegetation formations could be summarized in the main text or supplementary materials to give a clear idea of the sampling patterns.

Response: We agree that plot-size heterogeneity needs to be more clearly described and considered in the analyses associated with species richness (Figure 8). In our field survey, plot size was not designed to be uniform across all vegetation types, but was determined according to community physiognomy and stand heterogeneity. In general, the sizes of grassland plots were mainly 1 m × 1 m, of shrubland plots were mainly 5 m × 5 m, and of forest plots were mainly 10 m × 10 m or 10 m × 20 m. Local deviations occurred because of topographic constraints, especially slope, and field operability in complex mountain–valley terrain.

To make the sampling design more transparent, we have added a supplementary figure and table representing the spatial and elevational distributions of plot areas and their distributions among vegetation formations (Fig. S1; Table S2). These results show that the dominant plot areas were 1 m^2 for grasslands, 25 m^2 for shrublands, and 100 m^2 or 200 m^2 for forests, and that the major plot-area classes were broadly distributed across the surveyed region and elevational gradients.

We acknowledge that the original description of species richness along elevation did not explicitly account for plot-size heterogeneity. Therefore, we revised the analysis of elevational richness patterns. We did not use rarefaction, because abundance records were based on the number of individuals or clumps and are not directly comparable across vegetation plots. Instead, we analyzed plot-level species richness separately for forest, shrubland, and grassland plots using generalized additive models, with elevation as a smooth term and plot area included as a covariate. The fitted trends were then standardized to representative plot areas of 100 m^2 for forests, 25 m^2 for shrublands, and 1 m^2 for grasslands. Besides, the previous mountain-scale comparison was removed because sample number within different vegetation types were insufficient for robust modelling (Figure 8a).

In addition, the elevational patterns of growth-form and life-form composition were recalculated at the plot level (Figure 8b–d). Specifically, we first calculated the relative proportion of each growth-form or life-form category within each plot and then averaged these plot-level proportions within each elevational belt. It is helpful to reduce the influence of plot number and plot size among elevational belts. The corresponding Methods, Results, Figure 8, and caption have been revised accordingly.

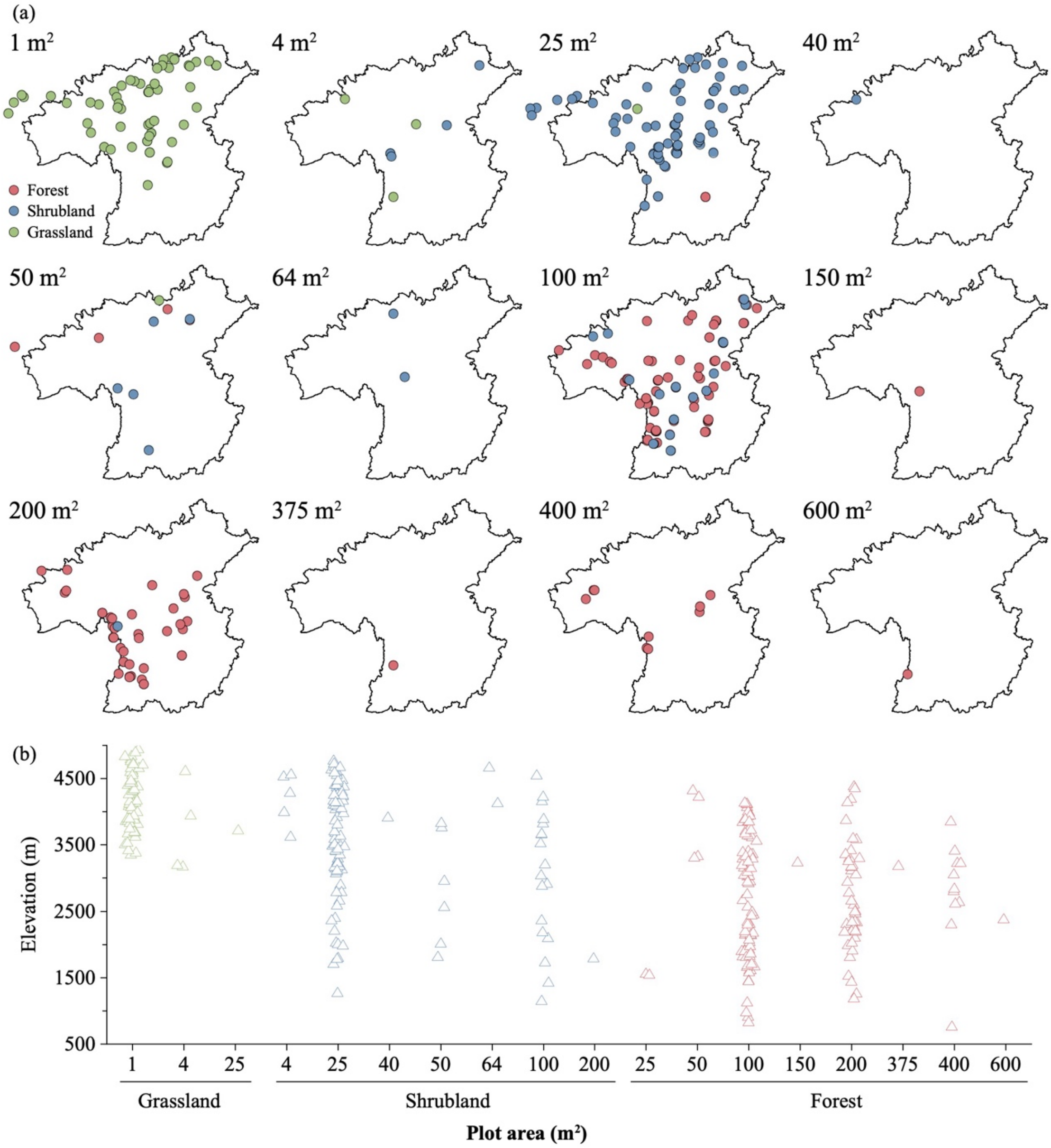

Figure S1. Horizontal **(a)** and elevational **(b)** distributions of plot areas in the HDM-Plot dataset.

Table S2. Distribution of plot areas among vegetation formations in the HDM-Plot dataset.

Area (m^2)	1	4	25	40	50	64	100	150	200	375	400	600
Forest												
DNF	/	/	/	/	/	/	3	/	1	/	/	/
DENF	/	/	/	/	/	/	/	/	1	/	/	/
ENF	/	/	/	/	1	/	28	/	13	1	7	/
NBF	/	/	/	/	/	/	8	/	12	/	/	/

DBF	/	/	/	/	2	/	18	/	6	/	1	/
EDBF	/	/	/	/	1	/	9	/	1	/	/	/
EBF	/	/	2	/	/	/	13	1	7	/	2	1
RF	/	/	/	/	/	/	1	/	1	/	1	/
Total	/	/	2	/	4	/	**80**	1	**42**	1	11	1
Percentage	/	/	1.4	/	2.8	/	**56.3**	0.7	**29.6**	0.7	7.7	0.7
Shrubland												
ENS	/	/	7	/	1	/	2	/	/	/	/	/
NBS	/	/	3	/	/	/	/	/	/	/	/	/
DBS	/	1	29	1	5	1	4	/	/	/	/	/
EDBS	/	/	11	/	/	/	2	/	1	/	/	/
EBS	/	4	25	/	/	1	**10**	/	/	/	/	/
BS	/	/	2	/	/	/	/	/	/	/	/	/
Total	/	5	**77**	1	6	2	18	/	1	/	/	/
Percentage	/	4.5	**70.0**	0.9	5.5	1.8	16.4	/	0.9	/	/	/
Grassland												
TG	34	1	/	/	/	/	/	/	/	/	/	/
RG	10	1	/	/	/	/	/	/	/	/	/	/
FG	13	1	/	/	/	/	/	/	/	/	/	/
SSG	/	1	1	/	/	/	/	/	/	/	/	/
Total	**57**	4	1	/	/	/	/	/	/	/	/	/
Percentage	**91.9**	6.5	1.6									

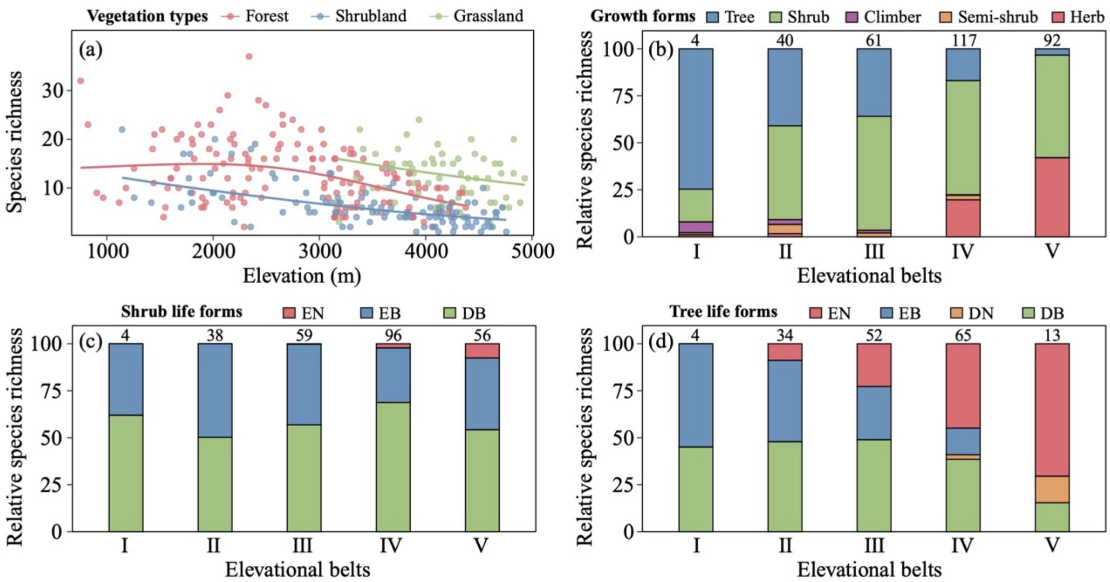

Figure 8. Elevational patterns of plot-level species richness **(a)**, growth forms **(b)**, and woody life forms **(c, d)** in the HDM-Plot dataset. In panel **(a)**, points represent observed species richness in individual plots, and fitted lines show elevational trends estimated using generalized additive models with plot area included as a covariate. Predictions were standardized to representative plot areas of 100 m^2 for forests, 25 m^2 for shrublands, and 1 m^2 for grasslands. Panels **(b–d)** show mean within-plot proportions across elevational belts. Elevational belts are defined as I, 0–1000 m; II, 1000–2000 m; III, 2000–3000 m; IV, 3000–4000 m; and V, 4000–5000 m. n denotes the number of plots included in each elevational belt. DB, deciduous broadleaf; DN, deciduous needleleaf; EB, evergreen broadleaf; and EN, evergreen needleleaf.

Comment 2 Strengthen the dataset description by summarizing more of the measured variables beyond species composition. The assessment of the dataset largely constrained to species diversity in its current form. Since the dataset includes structural and abundance-related information, a few additional summaries would help readers better appreciate its value and scope. For example, patterns of DBH, community height, coverage, or abundance would be informative, especially for forest and grassland plots. It may also be useful to show how species richness varied across disturbance levels so the readers would have a better understanding of the potential limitation of the dataset, given that disturbance intensity was recorded for each plot and many of the plots were sampled along the road.

Response: We agree that the original manuscript focused mainly on species diversity and composition, whereas the dataset also includes important community variables. To better demonstrate the scope and value of the dataset, we have expanded Section 4.4, Vegetation classification, by describing how these community variables vary among vegetation formations in relation to their horizontal and elevational distribution patterns. Accordingly, we have summarized species richness, community height, community coverage, and woody structural measurements such as diameter at breast height (DBH) and basal diameter (BD) in the newly added Table 3. In addition, we have examined the elevational patterns of these community variables, as well as the distribution of species richness among disturbance intensity. However, we did not present these exploratory analyses as major descriptive results of the dataset or use them for ecological inference, because they are strongly affected by vegetation type, plot number, plot size, and the non-experimental sampling design. Moreover, disturbance intensity was recorded as a relative plot-level descriptor rather than as part of a pre-designed disturbance gradient. The corresponding text and Table 3 have been added to the revised manuscript.

Table 3 Summary of community characteristics and structure in the HDM-Plot dataset

Formations	SR	Community height (m)	Community Coverage (%)	DBH or BD (cm)
DNF	10 ± 2	10.5 ± 3.9	71 ± 13	11.9 ± 5.1
DENF	4	15.6	35	10.1
ENF	12 ± 6	16.7 ± 7.4	66 ± 14	17.1 ± 11.6
NBF	14 ± 7	18.9 ± 6.4	68 ± 17	16.1 ± 7.9
DBF	14 ± 7	13.1 ± 5.5	76 ± 13	10.2 ± 5.8
EDBF	15 ± 6	14.8 ± 6. 4	73 ± 10	10.7 ± 5.5
EBF	12 ± 5	14.7 ± 8.8	74 ± 16	15.4 ± 8.8
RF	24 ± 7	35.3 ± 11.9	80 ± 10	23.3 ± 10.9
ENS	5 ± 2	2.5 ± 2.3	78 ± 18	3.4 ± 1.8
NBS	6 ± 1	2.4 ± 1.0	60 ± 33	2.6 ± 1.6
DBS	7 ± 4	2.7 ± 1.7	61 ± 21	1.6 ± 0.6
EDBS	10 ± 6	3.6 ± 3.1	62 ± 22	1. 9 ± 1.2
EBS	5 ± 4	2.5 ± 2.3	69 ± 17	1.6 ± 1.1
BS	5 ± 4	5.4 ± 2.9	68 ± 11	2.7 ± 1.1

TG	12 ± 4	0.061 ± 0.033	81 ± 15	/
RG	14 ± 5	0.108 ± 0.102	79 ± 25	/
FG	11 ± 6	0.079 ± 0.055	77 ± 16	/
SSG	5 ± 3	0.300 ± 0.283	50 ± 28	/

Values are summarized at the plot level within each vegetation formation and are shown as mean ± SD. SR denotes species richness per plot. DBH or BD represents the mean diameter at breast height or basal diameter of woody species within each plot.

Comment 3 The vegetation classification may benefit from a quantitative classification approach. In the vegetation classification section, 314 plots are divided into 142 alliance groups, and many groups contain only one plot based on Table 3. It raises the question of whether this level of detail is appropriate and whether all units need to be listed in the main table. The manuscript could also benefit from commenting on whether a quantitative classification approach, such as clustering based on species composition, might serve as a useful complement to the current descriptive framework.

Response: We agree that Table 3 was too detailed for the main text, especially because the revised vegetation classification system of China provides a fine hierarchical classification and many alliance groups or alliances are represented by only one plot. To improve readability, we have moved it from the main text to the Supplementary Material as Table S4. In the revised main text, we focus more on vegetation formation groups and vegetation formations.

We also agree that quantitative classification can provide a useful complement to the descriptive vegetation classification framework. However, the primary aim of our classification was to maintain consistency with the revised vegetation classification system of China and to support future use of the dataset in vegetation classification, mapping, and the *Vegegraphy of China*. This framework uses field-based community physiognomy, vertical structure, constructive species, and species importance values. These attributes are difficult to fully represent using a clustering approach based only on species composition, especially in a dataset that spans forests, shrublands, and grasslands across strong environmental gradients.

To address this comment, we conducted an exploratory two-way indicator species analysis (TWINSPAN) based on species composition. The first-level division broadly separated alpine grassland plots from non-alpine grassland plots, indicating that the main compositional contrast in the dataset is consistent with major vegetation physiognomic differences. However, finer divisions were less directly aligned with the field-based vegetation formations, partly because they were influenced by local species turnover, rare taxa, and uneven sampling among vegetation types. We have also added a column to the vegetation classification sheet in the updated dataset, providing the final group name assigned to each plot from the TWINSPAN. Therefore, we have added a comment in the revised manuscript noting that quantitative classification provides a useful complementary perspective, but that the field-based classification following Guo et al. (2020) remains the primary framework for this data paper.

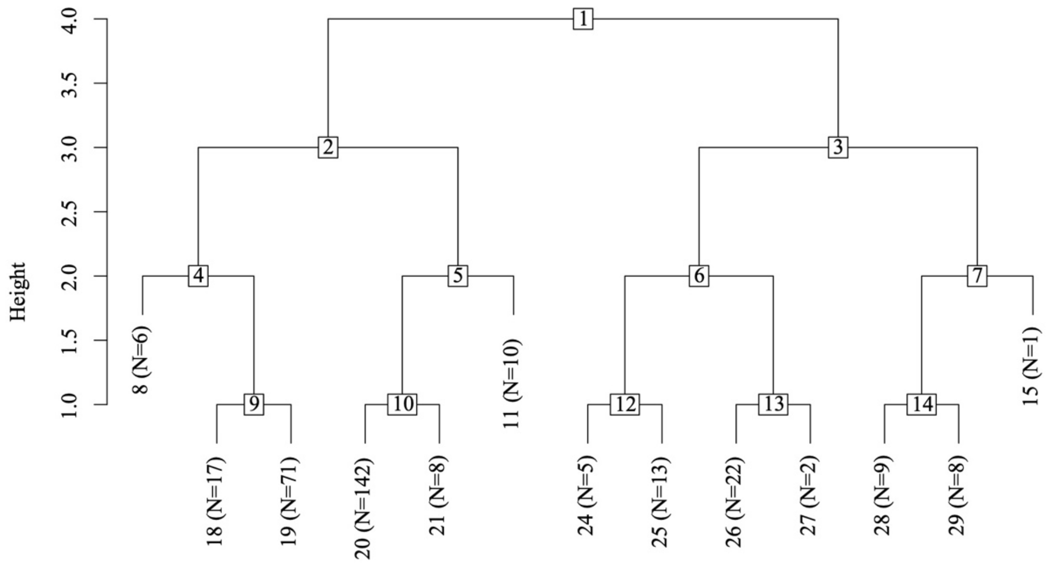

Figure S3 Two-way indicator species analysis of vegetation plots in the HDM-Plot dataset.

Comment 4 In addition, figures could be improved in regards of clarity. Fig. 2: Better to only highlight the few mountains that mentioned in the following analyses than numbering all the mountains in this region. Readers may feel it difficult to find where the five mountains used in Fig.8 for showing elevational gradient are. It would also be helpful to supplement the figure with plot density distribution against climatic space, as the deep valleys and mountains prevent the visualization of climatic condition change on the map and the dots representing single plot stacking on each other. Fig. 3-

6: Consider adding density plots or frequency maps to show the distribution patterns more clearly. This would make it easier for readers to assess how frequently each family or genus occurs across the surveyed plots.

Response: We have revised the figures to improve clarity and readability. In Figure 2a, we have removed the numbering of all mountains and rivers to reduce visual complexity. Because the mountain-scale analysis previously shown in Figure 8 has been removed from the revised manuscript, indicating that individual representative mountains is no longer necessary. As suggested, we have added a climatic-space coverage panel (Figure 2b), showing the distribution of forest, shrubland, and grassland plots along MAT and MAP gradients. In response to another reviewer's suggestion, we further added panels (c) and (d) to Figure 2 to compare the plot proportion with the relative area of each vegetation type and each elevational belt across the study region. Together, Figure 2 provides a clearer pattern of the realized sampling coverage in terms of geographical distribution, climatic space, vegetation types, and elevational belts. For Figures 3–6, the numbers in the lower-left corner of panels (a) and (b) represent the number and percentage of plots in which each family or genus occurs. This has been clarified in the updated figure captions. We further replaced the original jittered elevational scatterplots with violin plots with embedded boxplots in panels (c) and (d), which clearly show the elevational density and distribution range of dominant families and genera. A revised version of Figure 3 is shown here as an example, and the complete revised figures are included in the revised manuscript.

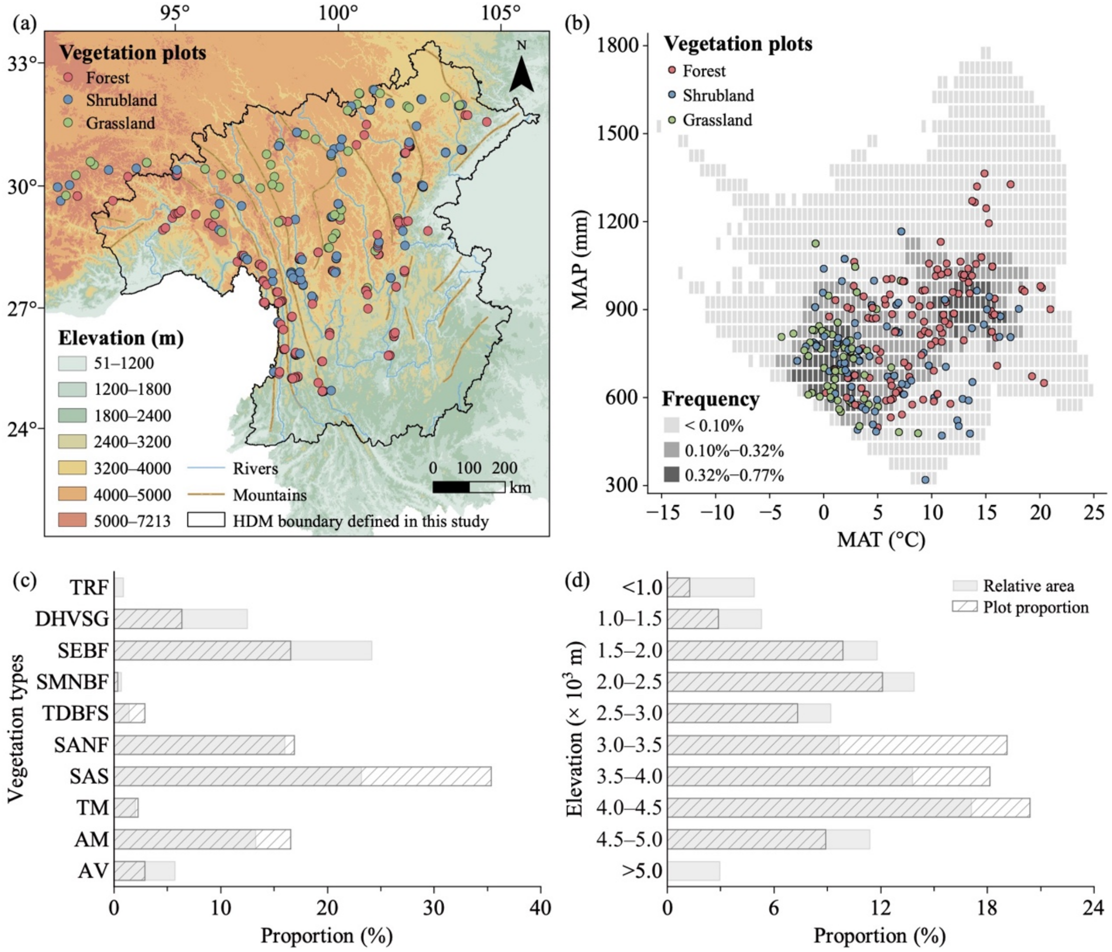

Figure 2. Spatial **(a)**, climatic **(b)**, vegetation-type **(c)**, and elevational **(d)** coverage of vegetation plots in the HDM-Plot dataset. Elevation data was derived from the SRTM 90 m dataset (Farr et al., 2007) and resampled into 1 km grid cells. Mountain and river data were obtained from the Digital Mountain Map of China Dataset (Nan et al., 2015) and Natural Earth (https://www.naturalearthdata.com, last access: 12 March 2026), respectively. Mean annual temperature (MAT, °C) and mean annual precipitation (MAP, mm) were derived from a 1 km monthly climate dataset for China covering 1991–2020 (Hu et al., 2025). Grey squares in panel **(b)** indicate the frequence of MAT–MAP combinations among all 1 km grid cells within the study boundary, based on two-dimensional bins of 0.5 °C for MAT and 50 mm for MAP. In panels (c) and (d), grey bars indicate the relative area of each vegetation type and each elevational belt, and hatched bars indicate the plot proportion of surveyed plots within each group. Vegetation types were extracted from the 1:1,000,000 Vegetation Map of the People's Republic of China (Editorial Committee of Vegetation Map of China, the Chinese Academy of Sciences, 2007a). TRF, tropical rain forest; DHVSG, dry-hot valley shrubby grassland; SEBF, subtropical evergreen broadleaf forest; SMNBF, subtropical mountains mixed needleleaf and broadleaf forest; TDBFS, temperate deciduous broadleaf forest and shrubland; SANF, subalpine needleleaf forest; SAS, subalpine shrubland; TM, temperate meadow; AM, alpine meadow; and AV, alpine cushion and sparse vegetation, and bare land.

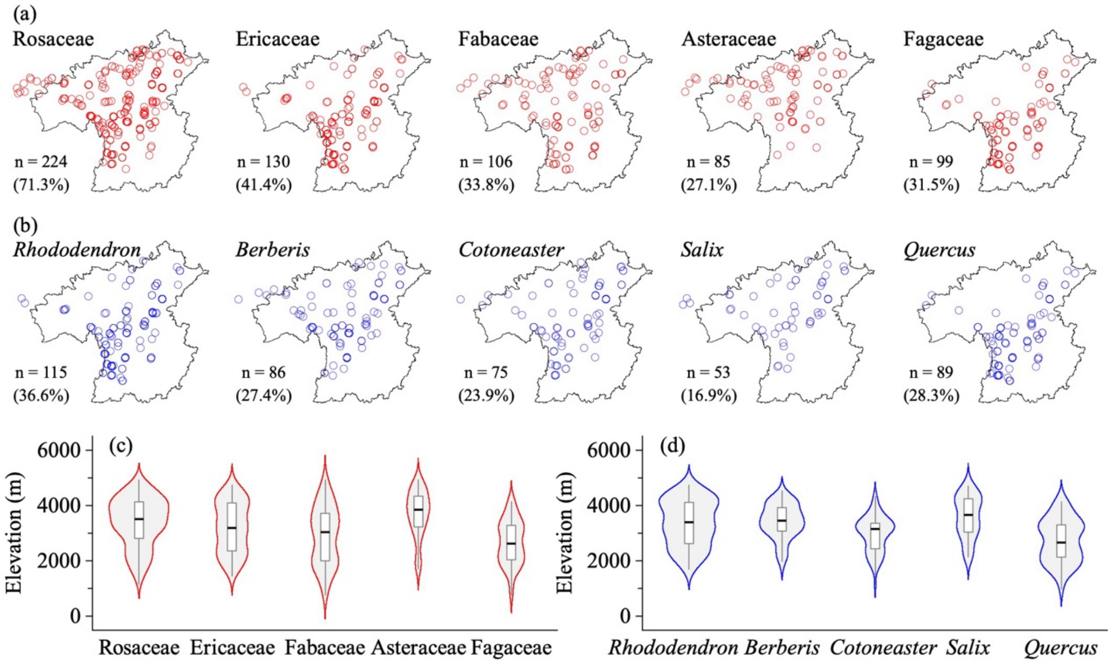

Figure 3. Horizontal **(a, b)** and elevational **(c, d)** patterns of dominant plant families **(a, c)** and genera **(b, d)** in all vegetation plots in the HDM-Plot dataset. n denotes the number of plots in which each dominant family or genus was recorded, and the values in parentheses indicate the proportion of the total survey plots.