Watershed unit is conductive to uncover microbial biogeography

Xiao-Yan Yang[1], wei Deng[1], Fa Zhang[1], Shuoran Liu[1], Davide Fornacca[1], Na Li[1], Yanpeng Li[1], Xi-Jie Fan[1], xin Zhang[1], Li-Lei Liu[1], Ya-Jiao Zhu[1], Kun Tan[1], Yin Gao[1], Wen Xiao[1], and Qinghua Cai[2]

[1]Dali University
[2]Institute of Hydrobiology Chinese Academy of Sciences

March 07, 2024

Abstract

Biogeography is flawed by a poor understanding of microbial distribution, due to the lack of systematical research framework, especially appropriate study units. We studied the biogeographic patterns of Nematode-Trapping Fungi by collecting and analyzing 2,250 specimens from 228 sites in Yunnan Province, China. We found typical watershed patterns both at the level of species and gene of Nematode-Trapping Fungi. The results showed that microbial biogeography could be better understood by 1) using watersheds as research units, 2) removing the coverup of widespread species, and 3) applying good sampling efforts and strategies. We suggest that watersheds could help unify the understanding of biogeographic patterns of animal, plant, and microbe, and may also help account for the historical and contemporary factors driving species distribution.

Introduction

Biogeographic patterns and the underlying mechanisms are essential in biodiversity research. The understanding of biogeographic patterns is crucial for predicting how global change affects biodiversity, designing effective conservation strategies, and achieving sustainable development (Gaston 2000). Studies on biogeographic patterns can be challenging due to the interactions among multiple biotic and abiotic factors, and to complex interrelationships among animals, plants, and microorganisms (Dolan 2006; Delavaux *et al.,* 2019). Previous research focusing solely on macrobiota biodiversity may therefore be inadequate to clarify the full view of biogeography (Mascarenhas *et al.* , 2020). Despite microbes being the most important biotic driver of global material and energy flows, biogeographic research on microbiota has long lagged behind that of macrobiota, and there are debates on whether heterogeneous biogeographic patterns exists in microbes (Martiny *et al.* , 2006; Meyer *et al.,*2018). The followings may be the main obstacles of microbial biogeography.

The primary obstacle for the lagged research on microbial biogeography is evaluating microbial diversity accurately. Microbial communities contain extremely high microbial biodiversities and nonculturable taxa. So, the traditional pure culture methods cannot deal with the obstacle. Recent advances in metagenomics have finally made fast evaluations of microbial diversity possible (Hanson *et al.* 2012), while concerns about the new methodologies remain. First, metagenomic approaches use operational taxonomic units (OTU) clustered based on arbitrary identity thresholds (usually 97%), which reduces the resolution of the taxonomic data set and generates confusing patterns (Storch *et al.* 2008). Second, metagenomic approaches are not concern on the strain level and unable to match OTUs with actual taxa or identify dormant microorganisms (Dolan 2006; Land *et al.,*2015; Locey *et al.,* 2020). Third, metagenomic approaches controversially neglect the rare taxa which are often more important than common taxa in analyses of microbial diversity (Gobet *et al.,* 2012). Although current reaserch have begun to use amplicon sequence variants (ASV) with higher resolution to replace OTU, ASV is still an operational taxonomic unit in essence, which has not been completely solved those problems.

At last, the core obstable is the choice of research units. Many studies about microbial biodiversity lack systematic sampling designs, clear research boundaries, and appropriate research scales (Ladau & Eloe-Fadrosh 2019). The worse is that biogeographic patterns are particularly sensitive to above factors (Bay *et al.,* 2020). Those caused confused results, and the confusions make it impossible for researchers to distinguish variation in microbial biodiversity caused by contemporary or historical factors (Castle *et al.,* 2019).

We propose three solutions to overcoming above obstacles and clarify the confusions. First, microbial biogeographic studies should depend on natural units with clear boundaries. Current studies on microbial biodiversity use either an arbitrary grid as unit of analysis or nothing, which may result in mixing up the effects of various factors. Natural units with clear boundaries should also be the fundamental need for the study about biodiversity, because coupling animal, plant and microbial biodiversity distributions is inevitable for future biogeographic research (Peters *et al.,* 2016). While biogeographic units for animals and plants have long been defined (e.g., ecoregions, biota, biogeographic provinces) (Hausdorf 2002; Smith *et al.,*2018), they do not match well with each other; great differences were among different regions and taxa (Smith *et al.,* 2020). We propose that natural watersheds are ideal units for biodiversity research, because they are relatively independent or "close" units of material and energy flow, and their boundaries can be natural barriers of transmission (Timur 2013). In addition, the contemporary and historical effects driving species distribution should change gradiently in different scales of watersheds, and watershed units could help us distinguish contemporary and historical effects.

Second, a framework that considers the importance of rare species and removing the coverup of widespread species should be established. Because rare species comprise a large part of the microbial community and contribute greatly to the microbial geographical patterns (Lynch & Neufeld 2015), their narrow distribution cannot be ignored. However, most of the existing research focuses on common species which widely distributed. The widespread species with large biomass are likely to masked the heterogeneous biogeographic patterns of microbial community. This may be due in part to the overconfidence in technology and resource limitations, so inadequate sampling efforts are engaged in current metagenomic approaches. Limitations on the resolution of metagenomic approaches often results in the removal of rare taxa data during data processing (Bay *et al.,* 2020). While the pure culture methods may be more effective at detecting rare species, its high workload makes it difficult to achieve comprehensive monitoring for huge microbial communities. To better account for the issue, we suggest 1) selecting a cultivatable and relatively small microbial taxon as the research object, 2) sampling adequately with stratified sampling strategy, 3) removing the coverup of widespread species at species level, and 4) analysing the widespread species biogeographic patterns at genetic level. These strategies will avoid the common problems plaguing the existing research system and should therefore set the standard for comprehensively understanding the biogeographic patterns of microorganisms.

Third, we suggest that studies about microbial biodiversity should be conducted in highly heterogenous environments. Microorganism dispersal is largely driven by both abiotic and biotic factors, and microbes are readily able to adapt and survive in a variety of environments. Indeed, studies conducted in homogeneous environments and within small spatial scales often reveal that either microorganisms are randomly distributed, or the weakness of environmental effects and dispersal limitations fail to indicate any non-random patterns in these region (Meyer *et al.,* 2018; Liu *et al.,* 2020). We believe it is easier to observe the spatial distribution pattern of microorganisms on large scales with high environmental heterogeneity.

Consequently, we conducted this research in Yunnan Province, China, which holds six major international rivers belonging to the Pacific and Indian Ocean systems. The watersheds were largely shaped by the collision of the Indian Ocean and Eurasian plates, with the formation of the Qinghai-Tibet Plateau. The spatial heterogeneity in Yunnan is very high, given that mountainous areas occupy 88.6% of the province.. Using the stratified sampling method, we collected 2,250 specimens from 228 sites in Yunnan Province (Figure 1, 2). We combined pure culture and molecular methods to investigate how common and rare Nematode-Trapping Fungi (NTF, a cultivable microbial group consisting of 99 known species) taxa are spatially distributed across Yunnan. By doing so, we aim to verify our three proposed recommendations regarding 1) using watersheds as the research units, 2) removing the coverup of widespread species, and 3) stratified sampling

from heterogenous environments, to overcome the challenges of microbial biodiversity research and to test our hypothesis that the microbial biogeographic patterns could be found in the perspective of watersheds.

1 Materials and Methods

1.1 Study area

Yunnan province is located between 21° 8' 32"˜29° 15' 8" N and 97° 31' 39"˜106° 11' 47" E, with a total area of 394,100 km^2 and spans an altitude of 76-6740 m. This area is characterized by sharp altitudinal gradient and the six large rivers – Mekong (澜沧江, Lancang Jiang), Irrawaddy (独龙江, Dulong Jiang), Salween (怒江, Nu Jiang), Red River (红河, Hong He), Pearl River (珠江, Zhu Jiang) and Yangtze River (金沙江, Jinsha Jiang) cutting across the province. This complex topography helped shape the region's relatively closed ecological zones, where biodiversity is extremely high.

1.2 Research objects

Nematode-Trapping Fungi (NTF) consist of eukaryotic microorganisms with terrestrial and aquatic biotypes belonging to the Orbiliaceae family. NTF are predatory microbes and, as their name suggests, prey on nematodes and other small organisms by producing trapping structures with their vegetative hypha. Globally distributed across various habitats and extreme environments, 101 NTF species have so far been reported, including those in the *Arthrobotrys* ,*Drechslerella* and *Dactylella* genera. All NTF species can be pure cultured and their morphological characteristics are observable under a low-magnification microscope. That NTF can be studied qualitatively and semi-quantitatively make them an ideal taxa for studying patterns in the spatial distribution of microbial diversity.

1.3 Stratified sampling

We systematically established 282 stratified sampling sites acording to Yunnan's hydrographic network. We were unable to sample from sites located in inaccessible terrain, including deep valleys and high mountains. Sites with strong human disturbance was also avoided. At last 228 randomly distributed sites (Ripley's K, $p = 0.2435$) were remained(Figure 2). To ensure the phenological consistency and clear aquatic-terrestrial boundary of the sites, sampling was conducted from the southernmost to the northernmost and between February 27th, 2014 and June 8th, 2014. At each site, 5 sampling points were set paralleled with the river both in the aquatic and terrestrial habitats with a distance of 10 meters (Figure 1b). Totally 2,250 sampling points were set (Figure 1a), 30 sampling points were lost due to no soil there. Using the five-point sampling method in 1m×1m squares (Figure 1b), samples consisted of ˜50 g of mixed soil (or sediment) and were collected at a depth of 0˜10 cm below the surface at each sampling point. Totally we collected 2,250 specimens from 228 sites. Samples were then stored in a disposable self-sealing bag at room temperature. We recorded habitat information, including location, longitude, latitude, and altitude of each sampling site. Samples were sent to the laboratory within 1 week of collection, immediately followed by NTF identification processing.

1.4 NTF classification and identification

Isolation and purification: Using the five-point spreading method, samples were plated on petri-dishes with CMA medium with ˜5,000 individuals of *Panagrellusredivivus* to induce NTF germination. We plated three replicates per sample for a total of 6,750 petri-dishes plated. The petri-dishes were incubated at room temperature (25-28). We examined the petri-dishes once after 3 weeks and again after 4-5 weeks using a stereomicroscope to isolate all the species. During each examination, we used single spore isolation to collect and purify any germinated NTF on CMA medium. A sterile toothpick was used during collection. Pure cultures were preserved until we were able to isolate and purify cultures from all specimens.

Morphological identification: All preserved pure cultures were revived in CMA medium and placed on a temporary slide using the inserting and sticking method. Morphological characteristics of NTF, including conidia, conidiophores and chlamydospore, were then photographed using a microscope camera (Olympus BX51, Japan). For each NTF strain, the type of trapping device was observed and confirmed using the observation chamber and nematode induction method. Strains were identified according to Li *et al.,* 2014.

Molecular identification: Each NTF strain was inoculated on PDA medium to gather the mycelium. We then extracted DNA from the mycelium using a CTAB protocol. The fragments of ITS (Internal Transcribed Spacer), TUB (β-tubulin gene), TEF (Translation Elongation Factor 1-alpha) and RPB2 (RNA polymerase II) were amplified by PCR and sequenced by BioSune Biotechnology Co., Ltd. (Shanghai, China). The homologous ITS and TUB sequences were aligned using the BLAST toolkit from NCBI (National Center for Biotechnology Information Search database).

We were able to identify and classify specific NTF species using morphological and molecular identifications, referring to the taxonomic system used by Li *et al.,* 2014.

1.5 Data analysis

1.5.1 Analysis of NTF species composition and distribution

The occurrence frequency (OF) for each species and sampling site was calculated respectively. OF (%) was calculated as: (number of specimens with NTF occurrence / total number of specimens) × 100%. Species with OF > 5% were defined as widespread species, while species with OF < 1% were defined as rare species.

Using the latitude and longitude of each site and NTF OFs, we then plotted distribution maps based on genus, species, and rare species with the *raster* , *sf* and *ggplot2* packages in R version 4.02.

The mean variance ratio was used for evaluating NTF species distributions. A mean variance ratio of 1 indicated a random distribution (conforming to the Poisson's distribution); a mean variance ratio >1 indicated a clustered distribution; a mean variance ratio <1, indicated an even distribution. The significance of these distributions was verified using a T-test.

We divided sampling areas into 9 grids using QGIS version 3.10. We selected the six grids with the largest sampling areas (grids A, B, C, D, E and F) for analysis (Figure S1 a). The watershed boundaries of Yunnan Province were extracted from a derived digital elevation model (CGIAR-CSI SRTM v4.1) at 90-meters resolution using GIS geoprocessing algorithms (Figure S1 b). We constructed upset plots using grid and watershed units with the *UpSetR* package in R version 4.02.

1.5.2 Genetic divergence and bigeographical distribution analysis of

Arthrobotrys oligospora

The DNA sequencing of ITS, TUB, TEF and RPB2 fragments was conducted on the strains of *Arthrobotrys oligospora* , which were found at every sampling site. The four sequences from each strain were summarized in a txt file and converted into a fasta file. MAFFT version 7 was used to generate the multi-sequence matrix (homologous sequences searching), and Bioedit was used to manually improve the accuracy of the alignment. We used jmodeltest software to select the optimal calculating alternative model for the piecing sequence. The phylogenetic tree was constructed following the Maximum Likelihood method (ML) in the way of partition calculation using IQ-Tree version 1.6.5 software. We used FigTree version 1.3.1, Microsoft Word (Microsoft office 2007) and Photoshop (Adobe Photoshop CS5 V12.0) to read the phylogenetic tree.

1.5.3 Biogeographical distribution of NTF in Yunnan

A first prediction was performed using the different watershed as unique predictors of the phylogenetic tree clades distribution of*Arthrobotrys oligospora* . We assigned clades to watersheds using the majority rule and assessed classification accuracy using error matrix metrics.

From the original dataset, we generated 45 different splits between train and test samples with train ratios comprising between 30% and 70% (step=5). Every ratio step included five different mixes of train/test samples. Voronoi (Thiessen) polygons were computed around each training set and assigned the clade corresponding to their originating point. The polygons were used to predict the test dataset clades.

1.5.4 Multivariate machine learning model

We compiled a dataset of 97 variables divided based on bioclimatic (24), topographic (6), vegetation (12), and soil properties (55) using QGIS (Table S1). In addition to the watershed, 19 predictors were selected among an initial pool of 97 variables. Table S2 shows the selection results after multinomial logistic regression and collinearity screening. In these cases, we always retained the values for the first layer.

To select the most suitable predictor variables to include in our machine learning model, we hierarchically screened each variable in our dataset. We iteratively preformed multinomial logistic regressions to assess the effects of each predictor variable on clade. We assessed the relationships among predictor variables using correlation matrices (Pearson's and Kendall's), and assessed multicollinearity using the Variance Inflation Factor.

The selected predictor variables and watershed predictors constituted the final dataset to investigate patterns in clade distribution using an Extra-Trees classifier (Extremely Randomized Trees). We aimed to improve the prediction ability from the previous analysis (watershed subdivision as a unique predictor), and therefore used the same accuracy metrics as in that analysis. Hyperparameters were optimized with Grid Search and Cross Validation techniques, though we also considered bootstrapping and accuracy estimations on the out-of-bag samples. Relative feature importance was calculated for the best model.

Results

The species composition and distribution of NTF in Yunnan Province

We identified 44 NTF species from 2,250 samples. These species belonged to three genera: *Arthrobotrys* (29 species), *Dactylella* (9 species), and *Drechslerella* (6 species) (Table S3). The rank order of occurrence frequency (OF) of NTF species corresponded to a logarithmic distribution model ($R^2 = 0.940$) (Figure 3). Four species were widely distributed (mean OF = 5.22%), and contributed to 9.09% of the total species richness and 62.32% of the total OF. Conversely, 31 species were classified as rare (mean OF = 0.18%), and contributed to 70.45% of the total species richness and 7.52% of the total OF.

A mean variance ratio of 1.07 ($t = 0.7029$,$p > 0.05$) indicated that NTF species were randomly distributed (Figure 4).

NTF spatial distribution patterns

At the genus level, *Arthrobotrys* was regionally distributed and found in 98.68% of the sampling sites, while *Dactylella* and*Drechslerella* were local distributed and found in 7.02% and 7.89% of the sampling sites respectively (Figure 5).

Among the 44 species detected in this study, the 4 widespread species were found in 83.30%, 56.58%, 47.37% and 41.23% of the sampling sites. The 31 rare species found in less than 10% of the sampling sites (range: 0.41-9.21%) (Figure 6).

When the all species were superposed in the map, NTF was found to be randomly distributed across the entire study area (mean variance ratio = 1.014, t = 0.153, p = 0.878). When the widespread species were removed, the cumulative distribution of rare and non-widely distributed species OFs was non-random and clustered (rare species: means variance ratio = 1.284, t = 3.022, p = 0.003; non-widely distributed species: means variance ratio = 1.326, t = 3.471, $p < 0.001$). The widespread species masked the non-random patterns.

NTF distribution: watersheds *vs* grids

Of the 19 rare species with non-single site distributions within grids or watersheds, 63.16% were distributed in the same or adjacent grid, and 84.21% were distributed in the same or adjacent watershed. With the exception of *Drechslerella dactyloides* , 8 of the 9 rare species that only appeared in 2 or 3 sites were distributed in the same or adjacent watershed (Figure 7).

At genetic level, the phylogenetic tree revealed that the widespread species *A. oligospora* was divided into 5 large clades. When grid was used as the unit of analysis, no clear distribution pattern of these clades

was detected. When watershed was used as the unit of analysis, the distribution of these clades was consistent with the natural watershed divisions in Yunnan Province, except for Irrawaddy River, which has no corresponding clade. The strains within the other clades were distributed within their corresponding watersheds. The 64.70%,60.71%,80.00%,63.89% and 61.11% of the strains from clade 1 to clade 5 were correspondingly distributed in the Yangtze River, Red River, Pearl River, Mekong, and Salween-Irrawaddy watersheds respectively (Figure 8).

Based on our phylogenetic tree, the spatial distribution *A. oligospora* was machine learned using the randomly generated Voronoi polygons and the watersheds. We found that the 45 maps generated using polygons all had lower accuracy than the maps watersheds generated (Table S4). On average, the accuracy of the polygons was low (mean: 36%, median: 38%), with all but one prediction falling below 50%. By assigning the clades according to the watershedes, 68.8% of clades were classified correctly (Figures S2, panel a). Watershed explained nearly 70 % of*A. oligospora* distribution in Yunnan. None of the additional climate, topographic, soil and vegetation variables significantly improved the model (Figure S3,Table S5).

We were better able to capture the distribution pattern of NTF when using watershed as the unit of analysis compared with grid (Figure 9). Only 17 intersections of were found when we constructed upset plots using watershed units. We found 19 species distributed in only one watershed unit, and that 76.2% of all species were distributed in the adjacent watersheds. There was a large gap between the dataset (variance was 7.89, and mean value was 2.53). Despite finding 22 combinations when constructing the integrative diagrams by grid, only 11 species were distributed in only one grid and 55.0% of all species were distributed in adjacent grids. There was a small fluctuation between the dataset (variance was 5.04, and mean value was 1.77). Ultimately, using watershed units led to a higher observation of endemic species, more species distributed within adjacent or similar units, and a more varied species composition, compared with using grid units.

Discussion

Widespread species masked thebiogeographic patterns of microbes

Nearly half of the reported NTF species (N = 44) in 3 genera were found during this study. At the species level, the distribution pattern was random and had no heterogeneous pattern. The number of common species only accounted for less than 5% of the total species but contributed to 70% of the OF. When excluding the 4 widespread species, the pattern tended to be non-random with a clear heterogenous pattern. If the distribution of the species with high OF were gradually superimposed, the patterns of NTF tended to be random. The results indicated that the extreme high OF and wide distribution of the common species mask microbial heterogeneous biogeographic patterns.

The 31 rare species were restricted to 0.44% ˜ 9.21% of the sampling sites. The genera *Dactylella* and*Drechslerella* were restricted in 7.02% and 7.89% of the sampling sites respectively. Because the limited distribution and the existence of endemic species were representative of the biogeographic pattern in itself[5], we concluded that NTF has heterogeneous biogeographic patterns..Although the rare species were low in abundance, they played a key role in understanding microbial biogeographic patterns. A low sampling effort may have limited any observations of these species. In this study, rare species (OF $<$ 1%) accounted for 70.45% of the 2,250 specimens collected in Yunnan Province. Had the sampling effort been reduced by 10%, we estimated that 29 of all species (93.54% of rare species) would have been more difficult to observe. Studies that used metagenomics to investigate microbial diversity patterns reported a sampling intensity interval between 3˜300 samples, with an average of 40 samples collected (summarised according to articles published in The ISME Journal and Microbiome over the past five years). Sampling efforts were subsequently reduced to same time and resources when using metagenomic methods (Karimi *et al* ., 2018). With low sampling efforts, many rare species would be ignored.

Watersheds are crucial for reveal microbial distribution patterns

Our results showed that the heterogeneous distribution of microbial diversity could be better explained by using watersheds as the units compared with grids. More endemic species were observed using watershed

units (19 species by using watersheds vs 11 species by using grids). Additionally, 84.21% rare species were clustered in the same or adjacent watersheds, while only 63.16% were clustered in the same or adjacent grids.

At the genetic level, the 5 clades of *A. oligospora* were consistent with the natural watershed divisions and these divisions were further supported by machine learning results. The 64.70%,60.71%,80.00%,63.89% and 61.11% of the strains from clade 1 to clade 5 were correspondingly distributed in the Yangtze River, Red River, Pearl River, Mekong, and Salween-Irrawaddy watersheds respectively. These findings suggest that watersheds are the most crucial factor in explaining the observed spatial distribution pattern of *A. oligospora* , more so than other environmental factors. Historical events leading to the formation of Yunnan's six large watersheds, including the uplift of Qinghai-Tibet Plateau and the spatial compression of the Hengduan Mountains, therefore likely shaped the spatial distribution of NTF genetic diversity. Because environments in mountainous region were geographically isolated by mountains, rivers, and other natural barriers, many new species evolved here (Ding *et al* ., 2020). As such, patterns of microbial diversity are easier to observe in this region given its clear watershed boundaries, complex terrains, high environmental heterogeneity, and abundant biodiversity.

Watersheds with clear boundaries and ecogeographical significance make them natural division units for biogeographic research. The dividing ridges between adjacent watersheds are critical dispersal barriers, the relatively isolated materials and energy flow were shaped. Previous studies have considered watersheds as ecosystem boundaries, and discussed the possibility of setting watersheds as the appropriate units in landscape ecology (Berkes *et al.,* 1998). For example, Paula et al. (2018) found that watersheds were better units when predicting diversity based on environmental heterogeneity. The hierarchical structure of watersheds also makes them appropriate for understanding how historical and contemporary factors contribute to diversity patterns. Moving from large scales (secondary or tertiary watersheds) to smaller scales (sub-watersheds), the influence of ridge barriers hindering distribution gradually decrease while environmental influences increase. The effects of historical events and contemporary environmental background are therefore perfectly coupled by watershed units. As such, using watersheds research units should help solve most of the problems in studying the spatial distribution of microbial diversity.

Meanwhile, from the perspective of watershed, the possibility of integrating plants, animals and microorganisms biogeographic research could be provided. As with microorganisms, biogeographic research on plants and animals is also plagued by confusing or deficient research units, methods, and analyses. For instance, the biogeographic provinces between the plants and animals were not coincident. Watershed units can provide an alternative that is applicable to multiple biological groups, allowing for research that consider the interactive effects of plants, animals, and microorganisms in shaping biogeographic patterns. For the core issues of biogeography, including the historical and environmental influences on the spatial distribution, the scale effects on the biological distribution patterns could be also solved. Finally, we believe that watersheds could support explanations on the origin of biodiversity, which is a fundamental scientific problem.

Conclusion

We did find definited biogeographic pattern of NTF in Yunnan. Heterogeneous pattern of NTF can not be found, if widespread species were included, or rare species were removed. Watershed unit, not grid, revealed clear distribution pattern of NTF at species and genetic level. Suggestions on future microbial biogeography studies include: employing watershed as basic unit; removing the impact of widespread species, valuing the contribution of rare species; employing stratified sampling strategy and good sampling effort; attaching importance to pure culture research and working on small taxa.

Acknowledgements: This work was funded by the National Natural Science Foundation of China (U1602262, 31360013, 31760126), supported by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP), Grant No. SQ2019QZKK2002. We thank Chen Jin-Hui (陈锦辉) from Kunming Institute of Zoology, Chinese Academy of Sciences for sample collection. Thanks to Jacob B Kraus, Ren Guo-Peng (任国鹏), She Rong (佘容), Zhang Cai-Cai (张彩彩) and Zhu Xiao-Ming (朱晓明) from Institute of Eastern-Himalaya Bodiversity research, Dali University for discussions. Thanks to the followings for their

contributions in laboratory research, they are Fang Yi-Hao (房以好), Zhou Han-Yu (周汉宇), Wang Jia-Liang (王家亮), Wang Chuan-Qi (王传琪), Zhou Xin-Juan (周新娟), Wang Hai-Qing (王海清), Ma Rui (马瑞), Li Fei-Teng (李飞腾), Li Jing-Chao (李静超), Jiang Lin (姜林), Gu Rong (顾荣), Wen Hui (温慧), Ni Jun (倪俊), Liao Yong-Hong (廖永红), Liu Sheng-Bin (刘胜斌), Duan Ji-Xia (段继霞), Liu Na (刘娜), Sun Hui-Fang (孙慧芳), Fan Xian-Qin (范先芹), Chen Xiao-Xi (陈晓溪), Wu Tao (吴涛), Mao Xu-Dan (毛旭丹), Wang Ting-Ping (王廷平), Lin Yan-Yun (林燕云), Huang Yuan-Tong (黄圆通), Xie Rong (谢蓉), Wang Yan-Fang (王彦芳), Xia Feng-Yu (夏丰雨), Guo Hao-Miao (郭浩妙), Zhang Ying (张莹), Li Yi-Lun (李宜伦), Yuan Jia-Xing (袁佳兴), Wang Shi-Xiong (王世雄), Zhou Fang-Yuan (周方圆), Xiong Meng-Jia (熊梦佳), Liu Tao (刘韬), Chen Zhi-Hong (陈志鸿), Lu Ya-Xian (陆亚娴), Gao Yu-Ting (高玉婷), Li Zheng-Qiang (李政强), Feng Xing-Xing (冯鑫鑫), Chen Yu-Xiu (陈玉秀), Zhang Xiao-Yu (张晓禹), Huang Zhi-Hao (黄志豪), Fan Ji-Zhuang (范集壮), Huang Fei-Hong (黄飞鸿), Ni Shui-Qing (倪水清), Zhu Zi-Zhong (朱自忠), Liao Guang-Hui (廖光辉), Wang Zhu-Hua (王柱华), Feng Bing (冯冰), Xu Xiao-Yu (徐潇雨), Yan Li-Hong (严丽红), Yue Xiao-Lan (岳晓兰), Teng Man (藤曼), Li Shuang-Mei (李双美), Zhao Jia-Jing (赵家静), Zhang Da-Bo (张大波), Li Zheng-Xiong (李正雄), Zhang Xun-Gang (张勋刚), Zhang Zan (张赞), Huo Bang-Guo (伙帮国), Yao Lu-Jun (姚路军), An Pan (安攀), Yang Zheng-Chao (杨正超), Yang Yao-Quan (杨耀权), Zhou Fa-Ping (周发平), Niu Kai-Yang (牛开阳), Xiong He-Ju(熊和菊), Lin Hui-Biao (林慧彪), Chen Xiao-Jun (陈晓俊), Zhou Jian-Fen (周建芬), Wang Ya-Juan (王娅娟), Qi Yu-Kun (齐宇坤), Huang Shuai-Yi (黄帅义), and Wang Ai-Di (王爱迪).

Author Contributions:Y.X.Y, X.W and C.Q.H conceived the study; Y.X.Y, X.W, D.W, Z.F, L.S.R and D.F undertook comparative and analytical work and wrote the paper; L.N, T.K and G.Y performed data analyses; Y.X.Y, X.W, D.W, Z.F, L.S.R, L.Y.P, F.X.J, Z.X, L.L.L, Z.Y.J coordinated the laboratory research work; Y.X.Y, X.W, D.W, Z.F, L.S.R, D.F, C.Q.H edited the manuscript.

Data accessibility statement: The data supporting the results will be archived in Dryad, and the data DOI will be included at the end of the article.

Competing interests: The authors declare no competing interests.

References

Bay, S. K., McGeoch, M. A., Gillor, O., Wieler, N., Palmer, D. J., Baker, D. J. et al. (2020). Soil Bacterial Communities Exhibit Strong Biogeographic Patterns at Fine Taxonomic Resolution. *mSystems* , 5, e00540-20.

Berkes, F., Kislalioglu, M., Folke, C., Gadgil, M. (1998). Exploring the Basic Ecological Unit: Ecosystem-like Concepts in Traditional Societies.*Ecosystems* , 1, 409-415.

Castle, S. C., Samac, D. A., Sadowsky, M. J., Rosen, C. J., Gutknecht, J., & Kinkel, L. L. (2019). Impacts of Sampling Design on Estimates of Microbial Community Diversity and Composition in Agricultural Soils.*Microb Ecol* , 78, 753–763.

Delavaux, C. S., Weigelt, P., Dawson, W., Duchicela, J., Essl, F., van Kleunen, M. et al. (2019). Mycorrhizal fungi influence global plant biogeography. *Nat Ecol Evol* , 3, 424–429.

Ding, W. N., Ree, R. H., Spicer, R. A., & Xing, Y. W. (2020). Ancient orogenic and monsoon-driven assembly of the world's richest temperate alpine flora. *Science* , 369, 578–581.

Dolan, J. R. (2006). Microbial Biogeography? *J Biogeogr* , 33, 199-200.

García-Llamas, P., Calvo, L., Cruz, M. D., Suárez-Seoane, S. (2018). Landscape heterogeneity as a surrogate of biodiversity in mountain systems: What is the most appropriate spatial analytical unit?*Ecological indicators* , 85, 285-294.

Gaston, K. J. (2000). Global patterns in biodiversity. *Nature* , 405, 220-7.

Gobet, A., Böer, S. I., Huse, S. M., van Beusekom, J. E., Quince, C., Sogin, M. L. et al. (2012). Diversity and dynamics of rare and of resident bacterial populations in coastal sands. *ISME J* , 6, 542–553.

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol* , 10, 497–506.

Hausdorf, B. (2002). Units in Biogeography. *Syst Biol* , 4, 648-652.

Karimi, B., Terrat, S., Dequiedt, S., Saby, N., Horrigue, W., Lelièvre, M. et al. (2018). Biogeography of soil bacteria and archaea across France. *Sci Ad* , 4, eaat1808.

Ladau, J., & Eloe-Fadrosh, E. A. (2019). Spatial, Temporal, and Phylogenetic Scales of Microbial Ecology. *Trends Microbiol* , 27, 662–669.

Land, M., Hauser, L., Jun, S. R., Nookaew, I., Leuze, M. R., Ahn, T. H. et al. (2015). Insights from 20 years of bacterial genome sequencing.*Funct Integr Genomics* , 15, 141–161.

Li, J., Hyde, K. D., Zhang, K. Q. (2014). *Nematode-Trapping Fungi* . Springer, Netherlands, pp. (13-40).

Liu, S, Wang, H, Tian, P, Yao, X, Sun, H, Wang, Q, Delgado-Baquerizo, M. (2020). Decoupled diversity patterns in bacteria and fungi across continental forest ecosystems. *Soil Biol Biochem* , 144, 107763.

Locey, K. J., Muscarella, M. E., Larsen, M. L., Bray, S. R., Jones, S. E., & Lennon, J. T. (2020). Dormancy dampens the microbial distance-decay relationship. *Philos Trans R Soc Lond B Biol Sci* , 375, 20190243.

Martiny, J. B., Bohannan, B. J., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L. et al. (2006). Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* , 4, 102–112.

Mascarenhas, R., Ruziska, F. M., Moreira, E. F., Campos, A. B., Loiola, M., Reis, K. et al. (2020). Integrating Computational Methods to Investigate the Macroecology of Microbiomes. *Front Genet* , 10, 1344.

Meyer, K. M., Memiaghe, H., Korte, L., Kenfack, D., Alonso, A., & Bohannan, B. (2018). Why do microbes exhibit weak biogeographic patterns?. *ISME J* , 12, 1404–1413.

Peters, M. K., Hemp, A., Appelhans, T., Behler, C., Classen, A., Detsch, F. et al. (2016). Predictors of elevational biodiversity gradients change from single taxa to the multi-taxa community level. *Nat Commun* , 7, 13736.

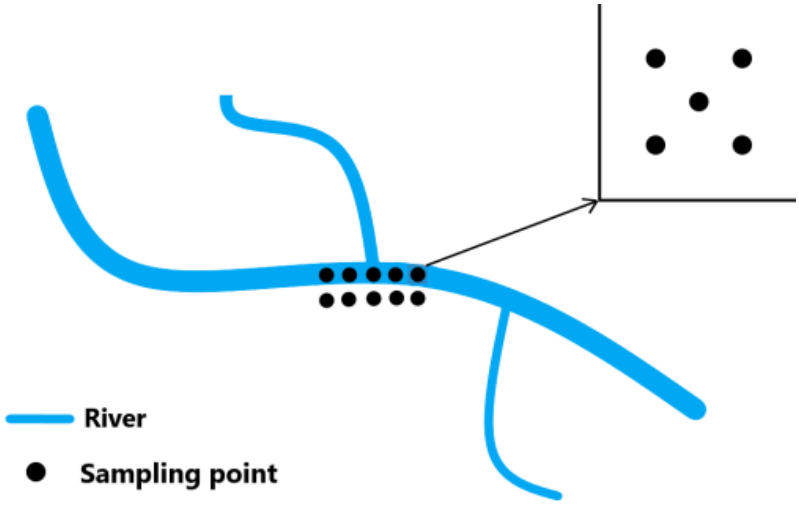
Smith, J. R., Hendershot, J. N., Nova, N., Daily, G. C. (2020). The biogeography of ecoregions: Descriptive power across regions and taxa.*J Biogeogr* , 47, 1413-1426.

Smith, JR, Letten, AD, Ke, PJ, Anderson, CB, Hendershot, JN, Dhami, MK. et al. A global test of ecoregions. *Nat Ecol Evol* , 2018 Dec;2(12):1889-1896.

Storch, D. & Izling, A. L. (2008). The Concept of Taxon Invariance in Ecology: Do Diversity Patterns Vary with Changes in Taxonomic Resolution? *Folia Geobot* , 43, 329-344.

Timur, O. (2013). *Advances in Landscape Architecture.* IntechOpen, London, pp. (176-206).

Ynch, M. D., & Neufeld, J. D. (2015). Ecology and exploration of the rare biosphere. *Nat Rev Microbiol* , 13, 217–229.



Figures

Figure 2 Geographic environments of the sampling sites.

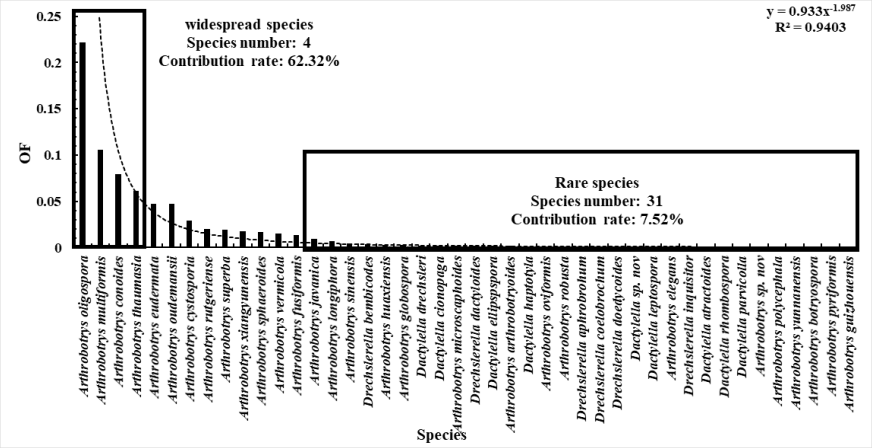


Figure 3 The species composition and occurrence frequency ordering of the Nematode-Trapping Fungi.

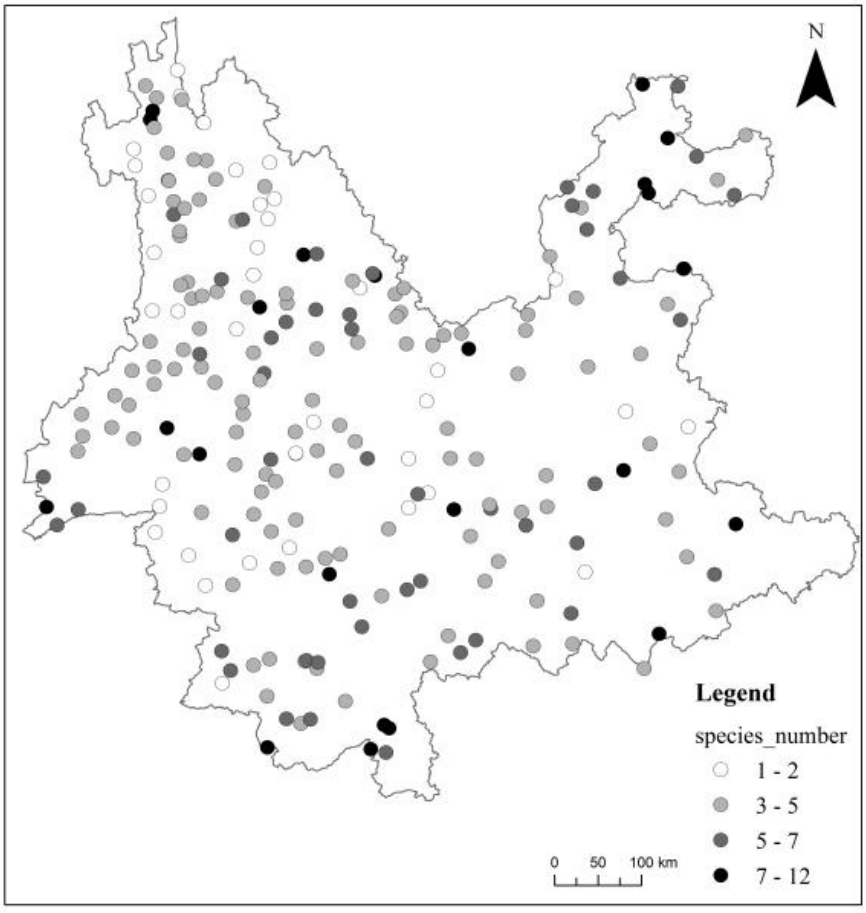


Figure 4 Biogeographic patterns of species number of Nematode-Trapping Fungi.

Figure 5 Biogeographic patterns of each genus of Nematode-Trapping Fungi.

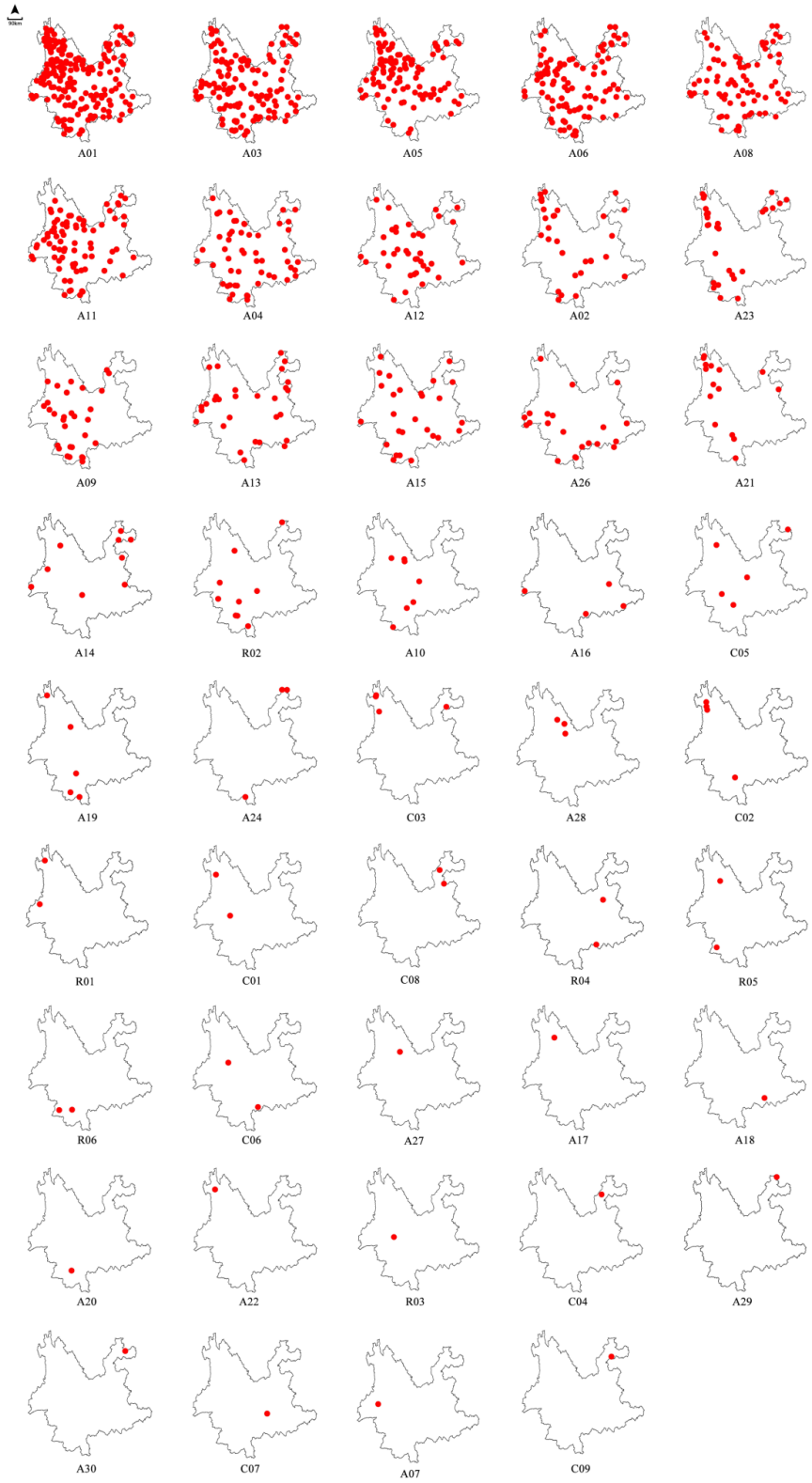
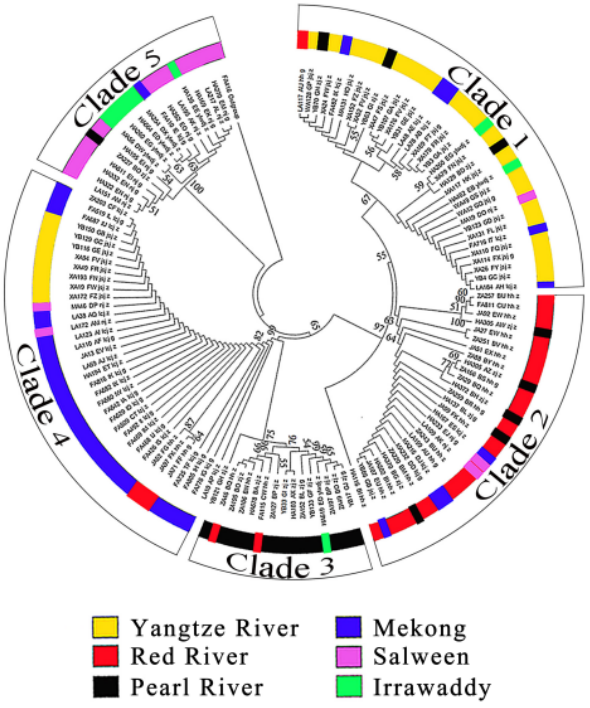
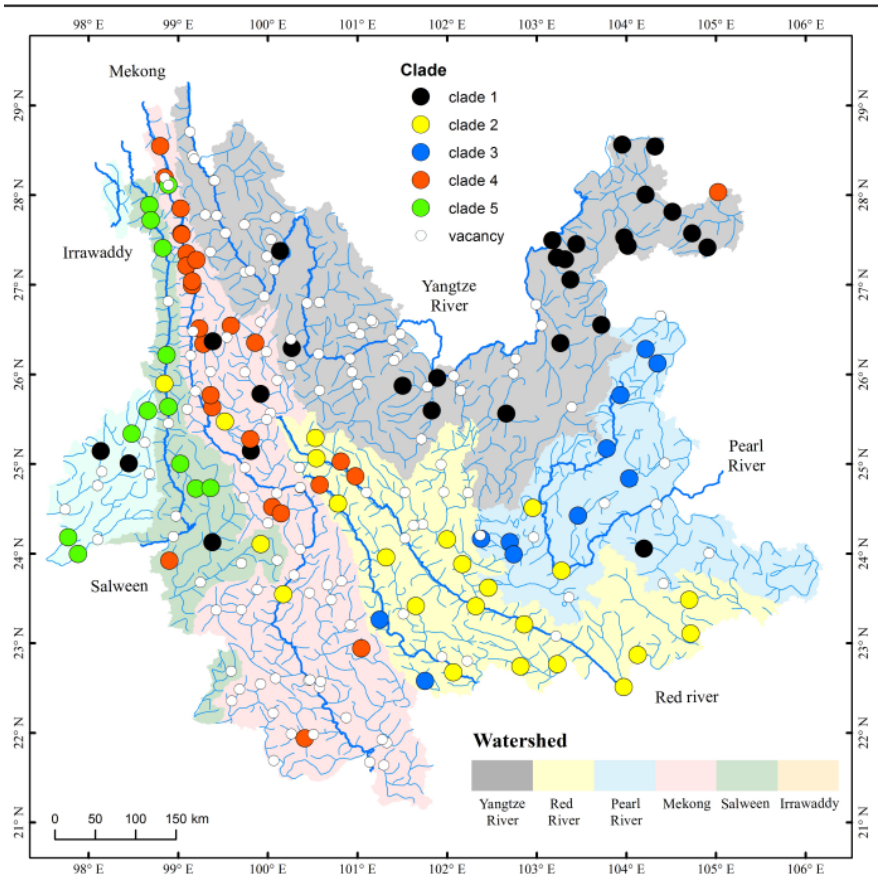


Figure 6 Biogeographic patterns of each species of Nematode-Trapping Fungi.

Figure 7 Biogeographic patterns of rare species in the pespective of a) grids and b) watersheds.





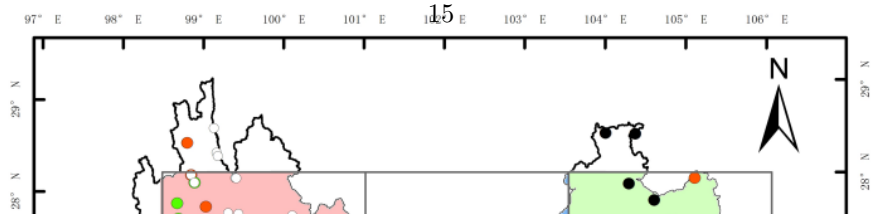


Figure 8 Biogeographic patterns of widespread species *A. oligospora* based on a) genetic clades respectively in the pespective of b) watersheds and c) grids

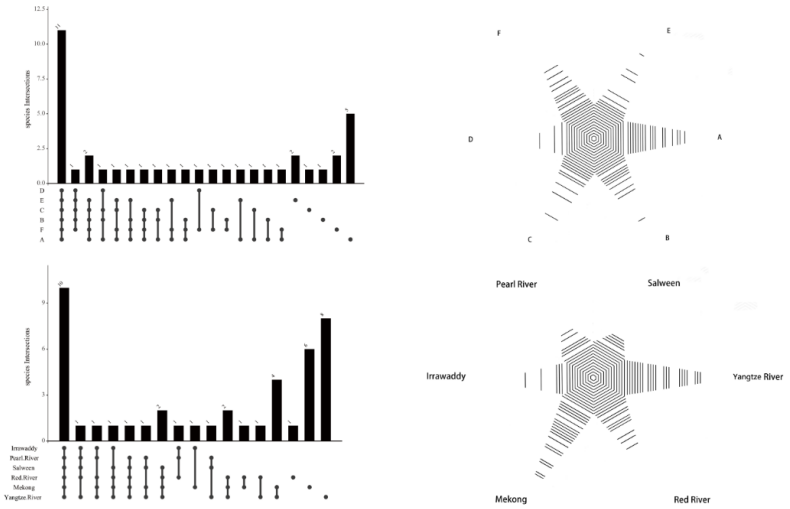


Figure 9 The upset plot of species in the perspective of a) grids and b) watershed units. The distribution pattern was uniform when using grids (c) and was heterogeneous when using watershed (d).