

1 Supplementary Materials 1

2 The User Guide of TaphonomeAnalyst 2.0

Sampling process



- 3
4 **Code availability:** The code of TaphonomeAnalyst 2.0 can be downloaded from
5 <https://github.com/wma1995/TaphonomeAnalyst2>
6 Development environments: Python 3.9.21 and R 4.1.3
7 **Article:** TaphonomeAnalyst 2.0: integrative analysis software of taphocoenosis co-occurrence
8 and geochemical data

9 Contents

10	Introduction	3
11	The basic idea behind the software.....	5
12	Data input	5
13	Assessment of sampling effort and estimation of theoretical maximum biodiversity	
14	(Module I)	8
15	Relative abundance of OTU analysis (Module II)	11
16	Proportion of taphonomic preservational grade of species analysis (Module III)	13
17	Taphonomic environment analysis (Module IV).....	17
18	Visualization of geochemical data (Module V).....	19
19	Assembling dissimilarity – environmental distance test (Module VI).....	21
20	Mantel Test between species abundance and ecological environmental variables (Module	
21	VII).....	22
22	Species correlation semi-matrix graphics (Module VIII)	24
23	Correlational Network Visualization (Module IX)	25
24	IX, 1 Overview	25
25	IX.2 Correlation.....	25
26	IX. 3 Louvain algorithm	27
27	Comparison of networks (Module X)	29
28	Frequently Asked Questions	34
29	The error of identical OTUs being identified as two separate ones.....	34
30	The proportion of taphonomic grades error	34
31	The network comparison error.....	34
32	The geochemical heat map error.....	34
33	Future developments	35
34	Acknowledgments	35
35	References.....	35
36	Developer.....	38
37		

38 Introduction

39 In recent years, there has been a significant increase in the number of studies aimed at
40 elucidating the structures and dynamics of ancient communities through taphonomic co-
41 occurrence networks (Guo Ma and Tang, 2023; Muscente, 2019; Xu, 2022). As taphonomic
42 co-occurrence data can reflect symbiotic relationships among various groups to some extent,
43 researchers are able to plot symbiotic networks by aggregating large amounts of taphonomic
44 co-occurrences. However, many of these studies have focused on large-scale marine biotas,
45 compared to smaller-scale lacustrine biotas that receive less attention (Muscente et al., 2019;
46 Xu et al., 2022). Marine research typically operates at large scales, allowing researchers to
47 deduce symbiotic (in the general and not mutualistic sense of the term) relationships from
48 presence/absence data alone (Muscente et al., 2019; Xu et al., 2022). In contrast, lacustrine
49 studies are conducted on a relatively small scale, where the transportation of fossil remains is
50 more complex and the sedimentary environment more variable. Obtaining co-occurrence data
51 for lacustrine research necessitates gathering of abundance data, which in turn requires field
52 excavation and lab identification of a large number of specimens, thereby rendering lacustrine
53 co-occurrence research difficult to implement over the long term.

54 Nevertheless, research on fossil lacustrine co-occurrence networks has also seen meaningful
55 progress. Our team has demonstrated that, even under less-than-ideal conditions such as
56 time averaging and varied transportation processes, a particular taphocoenosis still retains a
57 wealth of community-level information (Guo, Ma and Tang, 2023). We successfully mapped
58 the Daohugou faunal network from the late Middle Jurassic of China and divided it into
59 aquatic, edaphic, mudflat, and silvan modules (Guo, Ma and Tang, 2023). This exercise not
60 only provided statistically significant support for traditional networks based on morphological
61 function and taxonomic uniformitarianism, but also paves the way for further quantitative
62 studies in fossil community ecology.

63 Our team previously released the original version of TaphonomeAnalyst 1.0 (Guo, Ma and
64 Tang, 2023), designed for the study of small-scale, terrestrial, fossil assemblages. The
65 TaphonomeAnalyst 1.0 software package served as a comprehensive tool designed for the
66 downstream community analyses of taphocoenosis data, including abundances and
67 taphonomic preservational grades, primarily focusing on cluster analyses of communities and
68 community network analyses. TaphonomeAnalyst 1.0 integrated functions for importation,
69 analysis, and visualization of taphocoenosis data. The design idea of the software is based on
70 the accumulation of a substantial volume of OTU co-occurrence data from fossil sampling
71 plots that enable researchers to delineate species coexistence networks and discern various
72 environmental zones.

73 Although Taphonome Analyst 1.0 has core functions such as parsing ancient networks and
74 determining aquatic environments of entombed lacustrine communities, a principal limitation
75 of the original version of TaphonomeAnalyst 1.0 is its inability to explore the linkage between
76 ecological variables and community structure, which substantially affected its research value.
77 To address this constraint, we have developed an advanced iteration, TaphonomeAnalyst 2.0,
78 designed to expand the spectrum of ecological insights derivable from taphonomic data. This
79 updated version incorporates several enhancements:

- 80 (1) Integrates operational taxonomic unit (OTU) abundance with geochemical data for joint
 81 analysis.
 82 (2) Adds capacity for deducing synergies between biological differences and multiple
 83 geochemical factors.
 84 (3) Addi2 visualized co-occurrence networks from different environmental settings.

Module	Function	1.0	2.0
Assessment of sampling effort and estimation of theoretical maximum biodiversity (Module I)	S _{obs}	✓	✓
	Chao1	✓	✓
	Abundance-based Coverage Estimator (ACE)	✓	✓
Relative abundance of Operational Taxonomic Unit analysis (Module II)		✓	✓
Proportion of taphonomic preservational grade of species analysis (Module III)		✓	✓
Taphonomic environment analysis (Module IV)	Includes creation of Venn diagrams that compare the diversity found across sedimentary environments or outcrops.	✓	✓
Visualization of geochemical data (Module V)		✗	✓
Assembling dissimilarity environmental distance test (Module VI)		✗	✓
Mantel Test between species abundance and ecological environmental variables (Module VII)		✗	✓
Species correlation semi-matrix graphics (Module VIII)		✓	✓
Correlational Network Visualization (Module IX)	SparCC coefficient	✗	✓
	Pearson's coefficient	✓	✓
	Spearman's coefficient	✓	✓
	Kendall's rank coefficient	✓	✓
Comparison of networks (Module X)	Comparison of networks with different groups of sampling plots. Visualization of total nodes, total linked nodes, total edges, density, modularity, complexity, degree, and robustness.	✗	✓

85 **Table 1. Comparison of functions of the two TaphonomeAnalyst (1, 2) versions.**

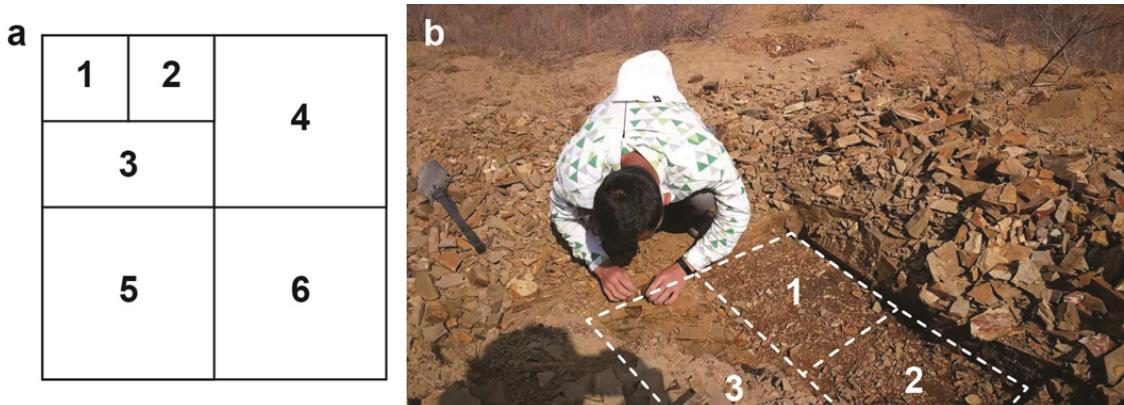
86 **The basic idea behind the software**

87 We deem it important to explain the rationale behind naming our software TaphonomeAnalyst
88 and to identify its research focus. Our work has been deeply inspired by the concept of the
89 “microbiome” (Berg et al., 2020; Dhariwal, 2017). Initially, the principal emphasis in microbial
90 ecology research was often placed on microbial communities, which are aggregations of
91 microorganisms that coexist in a shared environment (Berg et al., 2020). Due to the small size
92 of microorganisms and distinctive reproductive configurations, microorganisms can easily be
93 used in research settings by their various life-habit attributes. Consequently, microbial
94 communities are prone to hosting a multitude of transient visitors (Berg et al., 2020). These
95 itinerants typically occur at low abundances and lack ecological functionality. In 1988, Whipples
96 and his colleagues conceptualized the term "microbiome" as a fusion of "micro" and "biome"
97 and designated a "characteristic microbial community" within a "reasonably well-defined
98 habitat which has distinct physio-chemical properties" as its "theatre of activity" (Berg et al.,
99 2020; Whipples et al., 1988). The microbiome is an abstract concept derived from robust
100 statistical analyses of microbial community data at the technical level and the removal of
101 many accidental visitors. For a microbiome, also named an “abstract characteristic microbial
102 community”, the species co-occurrences and functions within such an assemblage notably
103 are discernible and can exhibit considerable responsiveness to shifts in environmental
104 conditions (Whipples et al., 1988). Paralleling the definition of a microbial community, the term
105 “taphocoenosis” encompasses an assemblage that includes a mixture of indigenous
106 organisms living in or near water and transient visitors transported from elsewhere. At this
107 juncture, we can successfully define the taphonome in ecology as characterized by fossil
108 communities in which component functions and co-occurrences can be observed and
109 distributed within a certain range of geochemical factors. The taphonome includes a wide
110 variety of species from aquatic to nearshore areas, with the possible exception of the largest
111 predators, whose fossilized remains typically are fragmented and scarce. In ecology, this
112 characteristic community has functional status; in stratigraphy, it has a stable geochemical
113 context within a consistent sedimentary environment.

114 **Data input**

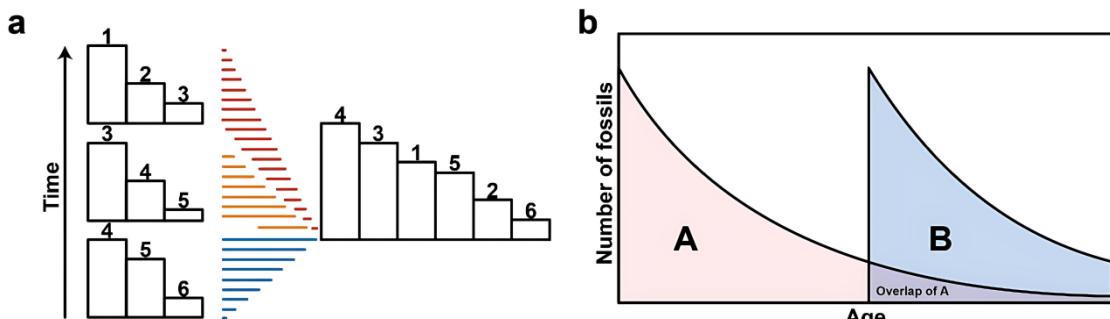
115 Fossil form a sample layer often from several centimeters of strata typically represent minimal
116 deposition spans, such as thousands of years, at least for most non-catastrophic depositional
117 environments. Fürsich and Aberhan (1990) used the term, time averaging, to describe the
118 minute chronological discrepancies existing within the same fossil layer that result from
119 bioturbation and the compression of strata. In paleoecological studies, establishing a definitive
120 temporal context for organismal burial remains a significant challenge. This limitation
121 becomes evident when considering that most fossil assemblages undergo temporal averaging
122 processes, incorporating remains from disparate communities or successive generations of
123 the same community over extended periods of time (Dhariwal et al., 2017; Karr & Clapham,
124 2015; Wing et al., 1992; Wright et al., 2003). Consequently, specimens recovered from
125 identical stratigraphic layers may represent organisms that perished and became interred at

126 distinct chronological junctures. The ecological interpretations of taphocoenoses formed
 127 through these variable depositional processes continue to provoke scholarly debate (Dhariwal
 128 et al., 2017; Karr & Clapham, 2015; Wing et al., 1992; Wright et al., 2003). Contemporary
 129 perspectives suggest that taphocoenoses may either reflect synecological communities under
 130 typical environmental regimes or represent diachronic assemblages accumulating within
 131 specific habitats over protracted time intervals (Dhariwal et al., 2017; Karr & Clapham, 2015;
 132 Wing et al., 1992; Wright et al., 2003).



133 **Nest sampling.** **a**, Schematic diagram of the nest sampling method. The numbers represent
 134 the steps of plots excavation. **b**, a fieldwork of nest sampling.

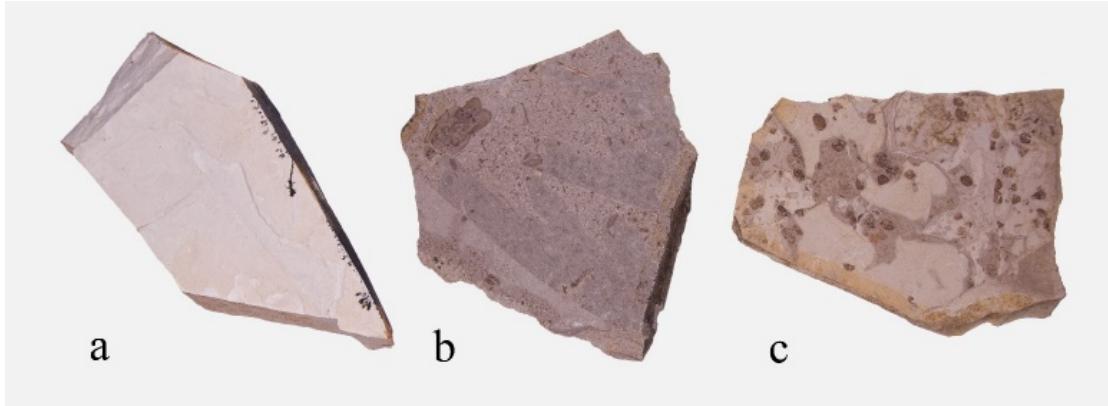
135



136
 137 **Impact of time-averaging on the relative abundance of species.** **a**, Time-averaging under
 138 a variable sedimentary environment (cited in Fürsich and Aberhan, 1990). The bars depict the
 139 numerical abundances of species 1 through 6. The fossil assemblages in the three different
 140 environments are mixed together after fossils have formed. Observations indicate that
 141 although the three assemblages become intermingled following time averaging, they exhibit
 142 negative quantitative correlations with one another. Statistical methods can be used to
 143 separate the three combinations. **b**, Overlap in time between two assemblage distributions
 144 (Thomas Olszewski, 1999).

145 TaphonomeAnalyst 2.0 draws its data from fossil tables and geochemical abundances in
 146 sampled layers. The fossil plots were excavated using nest sampling methods. All animal
 147 fossils were collected and identified. We suggest using operational taxonomic units (OTUs) at
 148 all taxonomic levels, instead of Linnean binomial names. An OTU typically represents a
 149 species-level taxon that is recognized and morphologically characterized but may not be
 150 formally described with a Linnean binomial. Such terminology streamlines ecological research
 151 by allowing analyses prior to the formal taxonomic description of the taxon in question. Using
 152 OTUs is a more annotated and convenient approach; for instance, Orthophlebiidae gen. et sp.

153 1. Some taxa, such as holometabolous insects involve immatures and adults of the same
154 species that often have large differences in values of indices characterizing their habitats and
155 morphological functions in the community. Users should mark immatures with "(I)" after the
156 scientific name to differentiate the immatures from the adults. The format of the sample
157 records is detailed in Supplementary Material 2.



158
159 **Typical lithologies of the fossil samples.** **a**, Lithologies of Beigou. **b**, Lithologies of
160 Donggou. **c**, Lithologies of Donggouli. Our fossil example sets were collected from the latest
161 Middle Jurassic Jiulongshan Formation near Daohugou Village, in Ningcheng County of Inner
162 Mongolia, China. Based on the stratigraphic interfingering of lacustrine and a volcaniclastic
163 apron facies, the sedimentary environment of Jiulongshan Formation is interpreted as several
164 fluctuating, shallow lakes with distant volcanogenic sedimentary input. Radiometric evidence
165 shows that the depositional age of the Jiulongshan Formation is 165–164 Ma and is time
166 equivalent to the Haifanggou Formation in Liaoning Province. All fossil samples should be
167 collected from the same stratigraphic layer without significant changes in the depositional
168 environment within, to ensure stable geochemical data, such as specimens a–c above.
169

170 The recommended output format is PDF, which is more suitable for editing vector elements.

171 Common parameters:

```
172 parent_parser = argparse.ArgumentParser(add_help=False)
173 parent_parser.add_argument('--input', type=str, required=True, help='Absolute or relative path
174 file.\t[e.g. "./data.xlsx"]')
175 parent_parser.add_argument('--format', type=str, default='pdf', choices=['png', 'svg', 'pdf'],
176 help='Output format.(default: %(default)s)')
177 parser = argparse.ArgumentParser(description='A comprehensive visual software for study
178 taphonome.')
179 parser.add_argument('-v', '--version', action='version', version='TaphonomeAnalyst 2.0')
180 subparsers = parser.add_subparsers(help='commands')
181
```

A	B	C	D	E	F	G
1 sample number	order	family	genera	species	taphonomic grade	
2 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
3 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
4 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
5 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
6 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
7 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
8 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
9 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
10 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
11 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
12 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
13 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
14 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
15 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
16 Dp001/0001	Conchostraca	Triglyptidae	Triglypta	Triglypta haifanggouensis	A	
17 Dn001/0001	Conchostraca	Triolventidae	Triolventa	Triolventa haifanggouensis	A	

182

183 **OTU identification form.** Each Sheet represents a sampled plot. Sample number refers to
 184 the unique identification number assigned to a fossil specimen. When multiple fossil
 185 individuals are present on a slab, each individual specimen should be documented separately
 186 by entering their details on each line. When identifying specimens, those that cannot be
 187 identified should be recorded as "unknown" in the corresponding biological classification
 188 category.

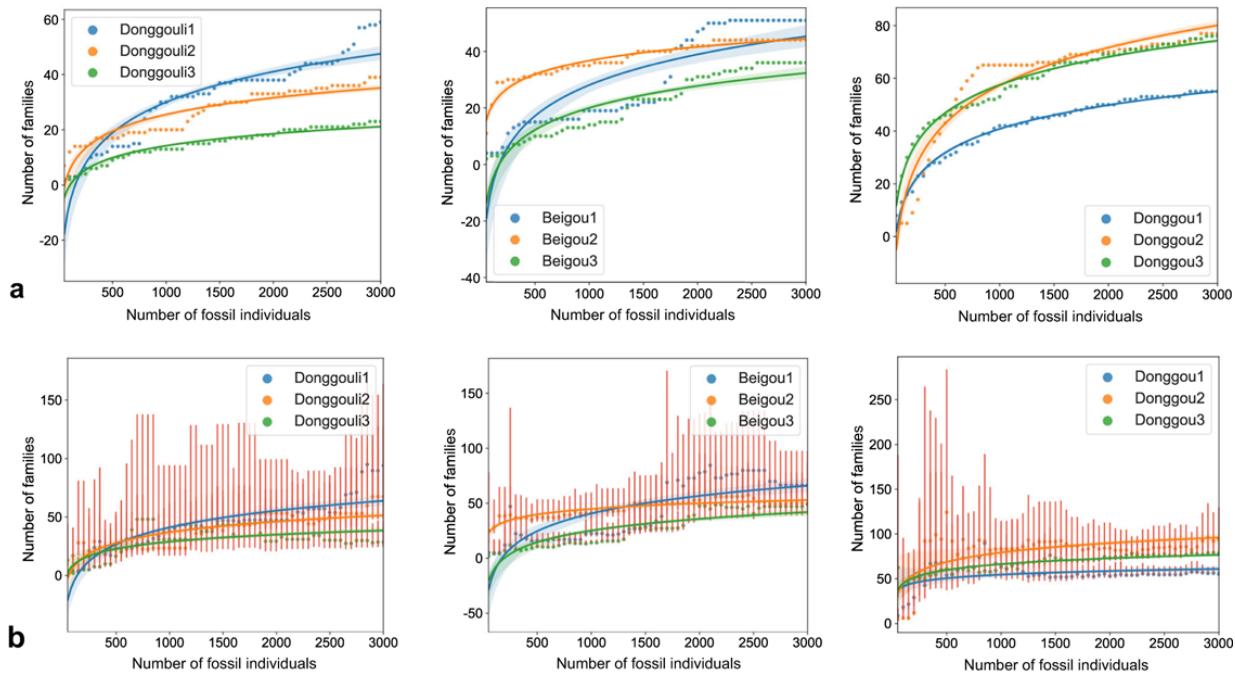
189

Assessment of sampling effort and estimation of theoretical maximum biodiversity (Module I)

191 This module has the capability to employ logarithmic curves for assessing sampling efforts
 192 and estimating the theoretical maximum biodiversity. It encompasses use of the S_{obs}
 193 (observed diversity), Chao1, and ACE (abundance coverage estimator) indices (Chao and
 194 Yang, 1993; Chao, 1984, 1992, 1993). In the field of lacustrine taphocoenosis research, it is
 195 often observed that the abundance of dominant aquatic species far exceeds that of terrestrial
 196 species, sometimes by several orders of magnitude. This disparity arises because terrestrial
 197 organisms had a much lower probability of fossilization when transported to a water body
 198 (Chao and Yang, 1993; Chao, 1984, 1992, 1993). Therefore, we recommend that users use
 199 Chao1 or ACE methods whenever possible to evaluate the sample coverage, as these
 200 methods are more sensitive to rare species.

201 S_{obs} , direct observational diversity, is suited to evaluate the coverage of sampling in strata
 202 where aquatic species do not have a significant advantage. During the sampling process,
 203 some Daohugou samples could contain over 2000 *Triglypta haifanggouensi* individuals
 204 among 3000 total individuals. In contrast, many terrestrial OTUs only have a few individuals.
 205 The S_{obs} curve may tend to flatten out when the number of samples is few, but sample
 206 location may still have significant role in potential diversity. Chao1 is sensitive to OTUs of only
 207 one individual, making it more suitable for plots where aquatic species dominate. The ACE

208 considers a wider range of rare species and makes corrections for the coefficient of variation
 209 and sample coverage, which is more reasonable. Nonetheless, due to the difference between
 210 the buried community and the present-day community, the definition of the abundance of rare
 211 species needs to be considered.



212 **Visualization of Sampling Coverage curves.** **a**, Sampling coverage curves (S_{obs}). **b**,
 213 Potential diversity curves calculated based on the Chao1 estimator. Users have the discretion
 214 to determine how many samples should be taken for a shift in each step of diversity. We fitted
 215 the sampling depth curve to the number of individuals based on the Chao1 and S_{obs}
 216 estimators for all sampling plots. The sampling curve calculates the diversity change of every
 217 50 samples and is fitted as a logarithmic function. The sampling depth-curve fit is in the form
 218 of a logarithmic function. The S_{obs} curve rises rapidly when the number of samples is less than
 219 500 and tends to be flat when the number of samples is more than 1000. When the sampling
 220 depth of all quadrats reaches about 3000 samples, the slopes of S_{obs} curve are
 221 0.002092–0.005316, which indicates that the sampling is sufficient. At the family level with the
 222 S_{obs} estimator at 3000 samples, the S_{obs} curve of Donggouli and Beiguou are lower, having
 223 only 21.1–47.5 families with the number of samples at 3000. In contrast, Donggou has 55.1 to
 224 80.1 families with the number of samples at 3000. The Chao1 curve was flat when there were
 225 about 500 samples. When the sampling depth of all quadrats reaches about 3000, the slopes
 226 of Chao1 curve are 0.001926–0.006907. At 3000 samples, Donggouli and Beiguou are also
 227 lower in Chao1 curve, with Bonggouli represented by only 41.9–66.0 families and Donggou
 228 with 61.0 to 96.0 families. The ratio between S_{obs} and Chao1 is 0.55 for Donggouli3, but for
 229 the rest it is 0.68–0.97.

230
 231 The Chao 1 estimator is:

$$\text{Chao1} = S_{obs} + \frac{F_1 (F_1 - 1)}{2(F_2 + 1)} \quad (1)$$

232 whereby S_{obs} is the direct observational diversity; F_1 is the number of OTUs whose
 233 abundance is one; and F_2 is the number of OTUs whose abundance is two.

235
 236 The ACE estimator is:
 237 $S_{ace} = S_{abund} + \frac{S_{rare}}{C_{ace}} + \frac{F_1}{C_{ace}} \gamma_{ace}^2$ (2)
 238 whereby S_{abund} is the count of abundant OTUs that typically includes those exceeding a
 239 rarity threshold—often set at 10 individuals. However, in our experience with fieldwork, a
 240 threshold of 10 individuals may be excessive. Our software allows the user to set the rare
 241 species threshold when all samples are pooled. S_{rare} is the number of rare OTUs (with less
 242 than or equal to rare threshold individuals) when all samples are pooled; C_{ace} is the sample
 243 abundance coverage estimator; F_1 is the number of OTUs whose abundance is one. γ_{ace}^2 , the
 244 estimated coefficient of variation for rare OTUs, is

$$245 \quad \gamma_{ace}^2 = \max \left[\frac{\sum_{i=1}^{10} i(i-1)F_i}{C_{ace}(N_{rare})(N_{rare}-1)} - 1, 0 \right] \quad (3)$$

246 Users have the flexibility to set the taxonomic level used for plotting sampling curves, with the
 247 default set at the family rank. Alternative options include order, family, genus, and species
 248 ranks.

249 **S_{obs}**

250 **Parameters:**

```
251 samplecurve_parser = subparsers.add_parser(name="m1sobs", parents=[parent_parser],  

252 help='sampling coverage curve. (Module I)\t[regplot]')  

253 samplecurve_parser.add_argument('--level', type=str, default='family', choices=['order',  

254 'family', 'genera', 'species'], help='Taxonomic level.(default: %(default)s)')  

255 samplecurve_parser.add_argument('--groups', type=str2dictlist, required=True,  

256 help='Grouping plots (Sheet names) with customized names.\t[e.g.  

257 "plotA:plotA1/plotA2,plotB:plotB1/plotB2"]')  

258 samplecurve_parser.add_argument('--output', type=str, default='./samplecurve',  

259 help='Absolute path or relative path and filename.(default: %(default)s)')  

260 samplecurve_parser.set_defaults(func=samplecurve)
```

261

262 **Command line:**

```
263 python ./TaphonomAnalyst2.py m1sobs --input "./Supplementary material2.xlsx" --level  

264 family --groups  

265 Donggou:Donggou1/Donggou2/Donggou3,Donggouli:Donggouli1/Donggouli2/Donggouli3,Bei  

266 gou:Beigou1/Beigou2/Beigou3
```

267

268 **Chao1**

269 **Parameters:**

```
270 chao_parser = subparsers.add_parser(name='m1chao', parents=[parent_parser],  

271 help='sampling coverage curve. (Module I)\t[regplot]')  

272 chao_parser.add_argument('--level', type=str, default='family', choices=['order', 'family',  

273 'genera', 'species'], help='Taxonomic level.(default: %(default)s)')  

274 chao_parser.add_argument('--groups', type=str2dictlist, required=True, help='Grouping plots (Sheet names) with customized  

275 names.\t[e.g. "plotA:plotA1/plotA2,plotB:plotB1/plotB2"]')
```

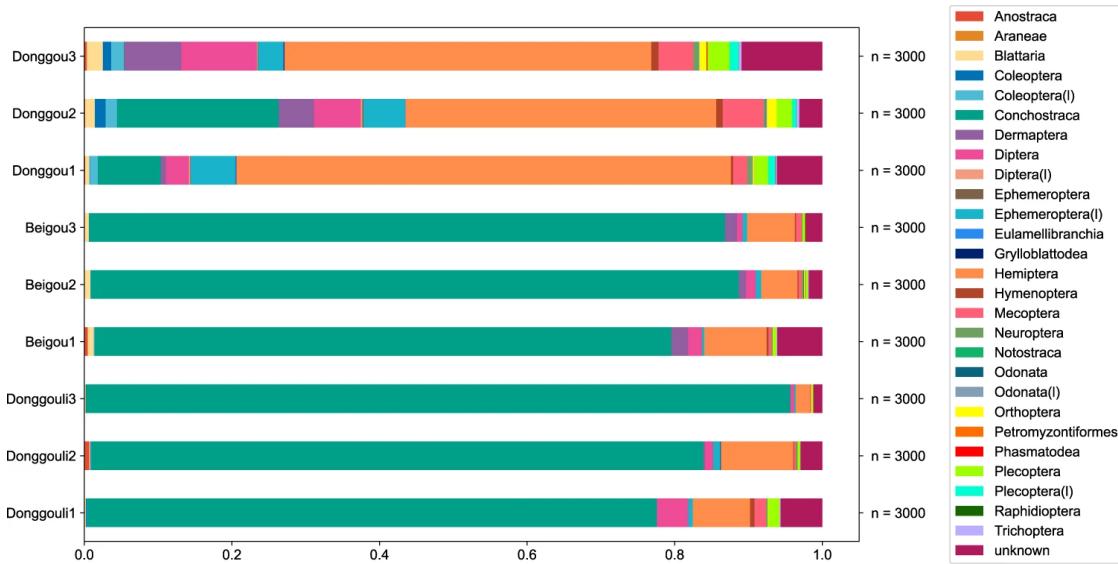
```

277 chao_parser.add_argument('--output', type=str, default='./chao', help='Absolute path or
278 relative path and filename.(default: %(default)s)')
279 chao_parser.set_defaults(func=chao)
280 Command line:
281 python ./TaphonomAnalyst2.py m1chao --input "./Supplementary material2.xlsx" --level
282 family --groups
283 Donggou:Donggou1/Donggou2/Donggou3,Donggouli:Donggouli1/Donggouli2/Donggouli3,Bi
284 gou:Beigou1/Beigou2/Beigou3
285
286 Ace: In the field of microbial ecology, the default abundance threshold for rare species is
287 set below ten. However, this value may be overestimated. It is recommended that users
288 set adjustments accordingly and provide explanations in their studies.
289 Parameters:
290 ace_parser = subparsers.add_parser(name='m1ace', parents=[parent_parser], help='ACE
291 potential diversity curve.(Module I)\t[regplot]')
292 ace_parser.add_argument('--level', type=str, default='family', choices=['order', 'family',
293 'genera', 'species'], help='Taxonomic level.(default: %(default)s)')
294 ace_parser.add_argument('--groups', type=str2dictlist, required=True, help='Grouping plots
295 (Sheet names) with customized names.\t[e.g. "plotA:plotA1/plotA2,plotB:plotB1/plotB2"]')
296 ace_parser.add_argument('--output', type=str, default='./ace', help='Absolute path or relative
297 path and filename.(default: %(default)s)')
298 ace_parser.add_argument('--rare', type=int, default=10, help='ACE rare threshold.(default:
299 %(default)s)')
300 ace_parser.set_defaults (func=ace)
301 Command line:
302 python ./TaphonomAnalyst2.py m1ace --input "./Supplementary material2.xlsx" --level family
303 --groups
304 Donggou:Donggou1/Donggou2/Donggou3,Donggouli:Donggouli1/Donggouli2/Donggouli3,Bi
305 gou:Beigou1/Beigou2/Beigou3 --rare 10

```

306 Relative abundance of OTU analysis (Module II)

307 This module facilitates the generation of bar graphs that illustrate species abundances. Given
308 that the abundance of fossils does not reliably indicate the actual diversity of the original
309 community, the relative abundance suggested by an OTU should be regarded as an imperfect
310 measure (McNamara et al., 2012; Smith and Moe-Hoffman, 2007; Wang et al., 2019). The
311 representation of species within a taphocoenosis, or fossil assemblage, is influenced by
312 factors such as the taxonomic unit investigated and its physical size (Smith and Moe-
313 Hoffman, 2007; McNamara et al., 2012). To some degree, these abundances can offer
314 insights into the varying source distances and trophic levels of the species present in the
315 fossil community.



316
 317 **Compositional proportion of the Yanliao Fauna by taxa from the sampled plots.** The
 318 taxon rank is the order. The sample sets of the three localities consist of 27,000 total fossil
 319 specimens, of which 25,796 (95.5%) specimens are identified to the taxonomic order level.
 320 The hydrophytic (aquatic) assemblage exhibited low diversity and high abundance. The three
 321 dominant hydrophytic species are the clam shrimp (conchostracan) *Triglypta haifanggouensis*
 322 (Triglyptidae), water boatman *Yanliaocorixa chinensis* (Corixidae), and mayfly *Mesobaetis*
 323 *sibirica* (Mesonetidae). *Triglypta haifanggouensis* occupies a very high percentage
 324 (77.2–95.4%) of the abundance from taphocoenoses at Donggouli and Beigou. In the
 325 Donggou plots, aquatic associations were dominated by *T. haifanggouensis*, *Y. chinensis*, and
 326 *M. sibirica*, accounting for, respectively, 0–21.9%, 34.8–63.2%, and 2.2–4% of the
 327 abundances. However, terrestrial assemblages often have unstable abundances and are
 328 elevated in diversity. Diptera (true flies), Mecoptera (scorpionflies) and Plecoptera (stoneflies)
 329 account for most of the abundance in plots where habitats are adjacent to water bodies,
 330 settings favorable for the formation of fossils. The numbers of Dermaptera (earwigs) and
 331 Blattaria (cockroaches) also are high, likely attributable to their toughened, leathery tissues
 332 that form the body surfaces and wings of these insects.

333
 334 Users have the flexibility to set the taxonomic level.
 335 **Parameters:**

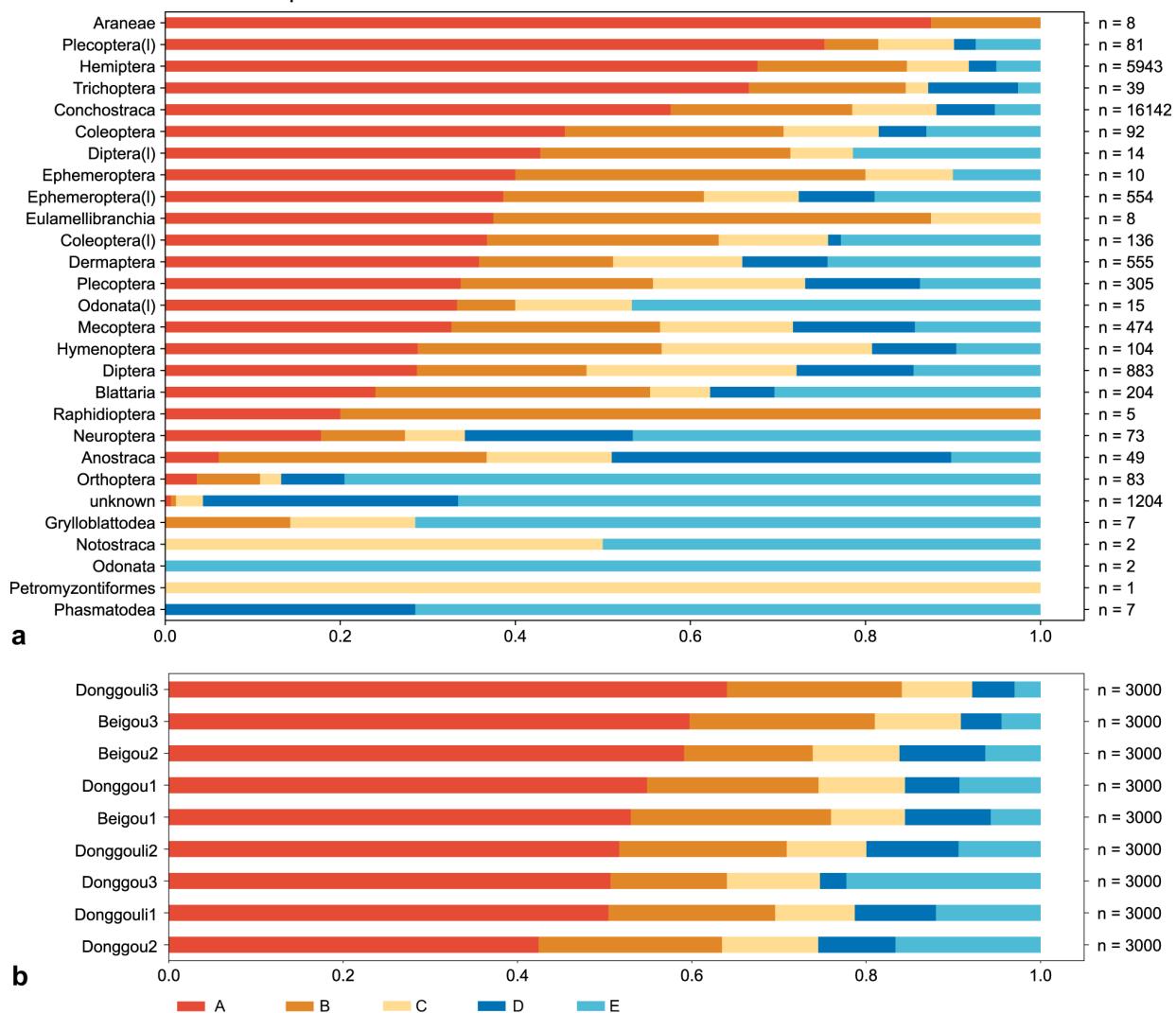
336 abundplots_parser = subparsers.add_parser(name='m2abundplots', parents=[parent_parser],
 337 help='Abundance-sampling plots. (Module II)\t[barh]')
 338 abundplots_parser.add_argument('--level', type=str, default='order', choices=['order', 'family',
 339 'genera', 'species'], help='Taxonomic level.(default: %(default)s)')
 340 abundplots_parser.add_argument('--output', type=str, default='./abundplots', help='Absolute
 341 path or relative path and filename.(default: %(default)s)')
 342 abundplots_parser.set_defaults(func=abundplots)

343 **Command line:**

344 python ./TaphonomAnalyst2.py m2abundplots --input "./Supplementary material2.xlsx" --level
 345 order

346 **Proportion of taphonomic preservational grade of**
 347 **species analysis (Module III)**

348 The taphonomic grade module evaluates the preservational quality of fossils, as reflected by
 349 their structural integrity and joint articulations (Guo, Ma and Tang, 2023). The taphonomic
 350 grade is categorized into five levels, ranging from A to E, where A represents the best
 351 preservation and E indicates the poorest (Guo, Ma and Tang, 2023). This module offers
 352 taphonomic grade bar graphs for various taxa. The taphonomic grade module assesses the
 353 preservational quality of the fossils based on their structural integrity, such as the extent of
 354 intact articulations of joints visible in the fossil. The taphonomic grade also can be used to
 355 interpret the distance between the original habitat and its eventual deposition into lake
 356 sediment. Although influenced by factors such as the robustness of body parts, particularly
 357 appendages, and body size, this method is extensively employed in taphonomic analyses.
 358 Users have the option to choose the level of classification for the taxa.



359
 360 **Proportion of taphonomic grades.** TaphomeAnalyst 2.0 offers the capability to selectively
 361 output the preservation levels of different OTUs (Operational Taxonomic Units) or the varying
 362 degrees of preservation across different sample plots. **a**, Proportion of taphonomic grades
 363 (A–E) by taxa. **b**, Proportion of taphonomic grades (A–E) by the sampling plot.

364
365 TaphonomicAnalyst 2.0 offers the capability to selectively output the preservation levels of
366 different OTUs (operational taxonomic units) or the varying degrees of preservation across
367 different sample plots. From an inventory of 27,000 individual fossils, we found that aquatic
368 organisms are well-preserved, mostly at a high level of body completeness. Although it cannot
369 be determined whether most *T. haifanggouensis* are articulated due to their burial position;
370 nevertheless, 78.4% are observed to have a complete chitinous shell and evident growth
371 bands. In addition, >87.8% of *Y. chinensis* are well-preserved, displaying entire bodies and
372 swimmeret appendages while 79.6% of mayfly nymphs have preserved gills and cerci.
373 Furthermore, beetles that have hardened wing covers, true bugs with leathery tegmina, and
374 cockroaches and earwigs that have highly keratinized bodies also are preferentially
375 preserved. By contrast, insects with softer body surfaces and more delicate wings, such as
376 nematoceran flies, stoneflies, mayflies, and katydids, are less well-preserved. Generally, the
377 distance between the insect-occupying habitat and a water body of eventual burial is one of
378 several decisive factors in the preservation of such insects. Trichoptera adults (84.6%),
379 Ephemeroptera adults (80.0%), and Plecoptera adults (48.1%), whose larvae and naiads live
380 in water bodies, have preservation grades of A or B. Mecoptera, Hymenoptera (sawflies and
381 wasps), Diptera, and Neuroptera (lacewings), which live in humid environments, have 56.5%,
382 56.7%, 48.1%, and 27.3% scores, respectively, at A or B grades. By contrast, Orthoptera
383 (grasshoppers and katydids), Phasmatodea (stick insects), and Grylloblattodea (ice crawlers)
384 have the least degree of preservation, consisting of 0–10% scores at A or B grades. By
385 comparing the taphonomic grades from the nine plots, we found that the differences were
386 small. About 70% of the individuals are preserved at the A and B grades. These observations
387 show that most fossils are buried in an autochthonous or sub-autochthonous manner, which is
388 an ideal context for sampling.

389
390 Proportion of taphonomic grades (A–E) of the taxa. Users need to set the taxonomic level
391 themselves.

392 **Parameters:**

```
393 TGotus_parser = subparsers.add_parser(name='m3otus', parents=[parent_parser],  
394 help='Taphonomic grades-taxa. (Module III)\t[barh]')  
395 TGotus_parser.add_argument('--level', type=str, default='order', choices=['order', 'family',  
396 'genera', 'species'], help='Taxonomic level.(default: %(default)s)')  
397 TGotus_parser.add_argument('--output', type=str, default='./TGotus', help='Absolute path or  
398 relative path and filename.(default: %(default)s)')  
399 TGotus_parser.set_defaults(func=TGotus)
```

400

401 **Command line:**

```
402 python ./TaphonomicAnalyst2.py m3otus --input "./Supplementary material2.xlsx" --level order  
403 Users need to set the taxonomic level themselves.
```

404 **Parameters:**

```
405 TGplots_parser = subparsers.add_parser(name='m3plots', parents=[parent_parser],  
406 help='Taphonomic grades-sampling plots (in customized order). (Module III)\t[barh]')  
407 TGplots_parser.add_argument('--groups', type=str2list, default=None, help='Environment
```

```
408     groups.(Recommend to group the plots by different aquatic and terrestrial environments)\t[e.g.  
409     "plotA1/plotB2,plotB1/plotA2,plotC1/plotC2"]'  
410     TGplots_parser.add_argument('--output', type=str, default='./TGplots', help='Absolute path or  
411     relative path and filename.(default: %(default)s)'  
412     TGplots_parser.set_defaults(func=TGplots)  
413  
414     Command line:  
415     Default sort  
416     python ./TaphonomAnalyst2.py m3plots --input "./Supplementary material2.xlsx"  
417  
418     Custom sort  
419     python ./TaphonomAnalyst2.py m3plots --input "./Supplementary material2.xlsx" --groups  
420     Donggou1/Donggou2/Donggou3,Donggouli1/Donggouli2/Donggouli3/Beigou1/Beigou2/Beigo  
421     u3
```

The standard of taphonomic grades	
A	 A fossil specimen of an insect, likely a mayfly or dragonfly, showing a complete body with head, thorax, and abdomen. A vertical scale bar is present.
B	 A fossil specimen showing a large, elongated shell fragment. A vertical scale bar is present.
C	 A fossil specimen showing a partially articulated body, possibly a beetle or fly, with some legs and wings preserved. A vertical scale bar is present.
D	 A fossil specimen showing several isolated leg fragments. A vertical scale bar is present.
E	 <p>Scattered bones.</p> <p>Scattered bones.</p> <p>Scattered bones.</p> <p>Scattered bones.</p>

Table 2. Definition of taphonomic grades. The taphonomic grade categorizes the preservational quality of fossils into five levels, ranging

423 **Taphonomic environment analysis (Module IV)**

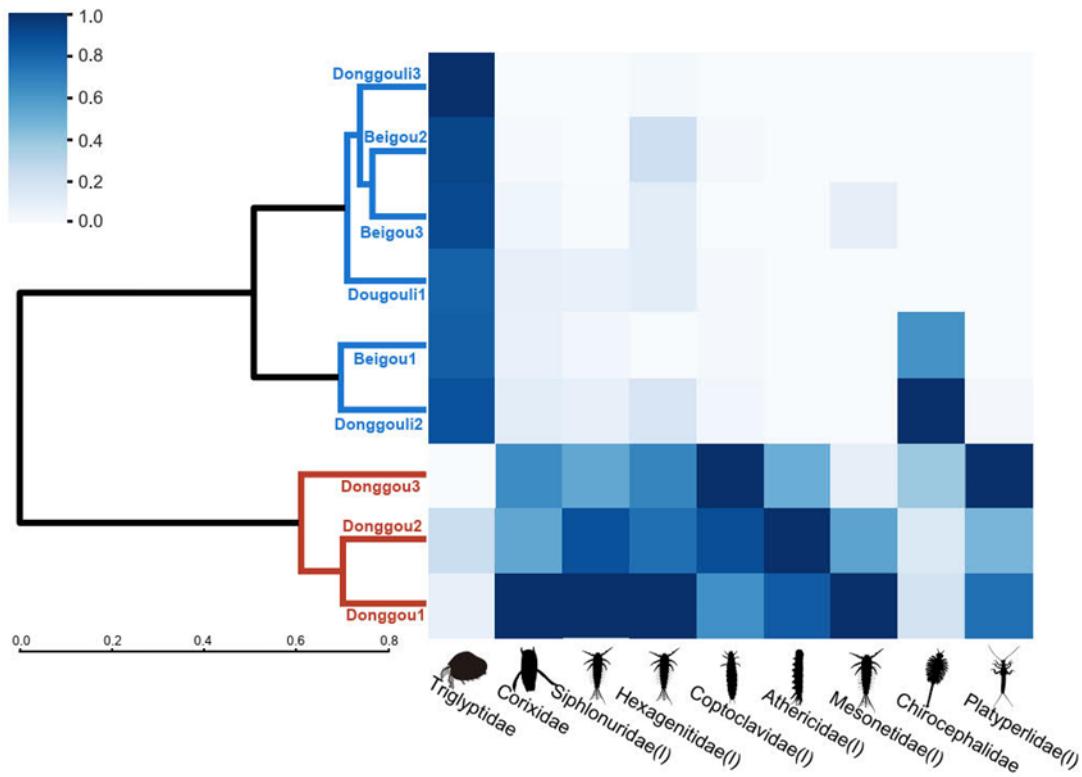
424 The primary aim of this module is to contrast and display the variations in species abundance
 425 across different sampling plots, and subsequent categorization of these samples into distinct
 426 groups. Since aquatic organisms typically undergo minimal transport during fossilization, their
 427 populations remain relatively stable, offering a robust reflection of the environmental
 428 conditions at the time of fossil deposition. The hierarchical clustering is calculated after a filter
 429 threshold is set at an individual count greater than 5. Hierarchical clustering is performed
 430 using the average linkage clustering method and Bray-Curtis distance metric, which are
 431 commonly employed in biodiversity studies. We suggest that users conduct a joint analysis by
 432 integrating environmental clustering, species distribution, and geochemical heatmap collages.
 433 By integrating geochemical heat maps, users can clearly discern the distribution of aquatic
 434 OTU abundances across different clusters, as well as associated differences in environmental
 435 factors. Additionally, the module generates Venn diagrams that illustrate the differences in
 436 diversity within different taphonomic environments.

437 The average assigns clustering method is:

$$d_{(u, v)} = \frac{d(u[i], v[j])}{\sum_{ij}(|u_i| * |v_j|)} \quad (4)$$

439 and the Bray-Curtis distance is:

$$d_{(u, v)} = \frac{\sum_i |u_i - v_i|}{\sum_i |u_i + v_i|} \quad (5)$$



441
442

443 **Diversity and aquatic abundance comparisons of different plots and sedimentary**
 444 **environments.** Hierarchical clustering is shown of nine sampling plots of sedimentary
 445 environments. The plots were clustered based on aquatic taxonomic abundance ($n > 5$) using

446 the average assigns clustering method and Bray-Curtis distance metric. The color of the heat
447 map indicates the normalized abundance of biological distribution. The clustering results show
448 that the sedimentary environments can be divided into two types.

449 **Aquatic species in the example dataset include:** *Daohugouectes primitinus* (I), *Triglypta*
450 *haifanggouensis*, *Triglypta haifanggouensis*, *Yanliaocorixa chinensis*, *Karataviella popovi*,
451 *Samarura gigantea* (I), Anisoptera fam. gen. sp. 1 (I), *Platyperla platypoda* (I), *Ferganoconcha*
452 *sibirica*, *Qiyia jurassica* (I), Mesomyzon sp. 1, *Triops* sp. 1, Chirocephalidae gen. sp. 1,
453 *Eurythoracalis mirabilis* (I), *Shantous lacustris* (I), *Foliomimus latus* (I), *Furvoneta viriosus* (I),
454 *Furvoneta raukus* (I), *Mesobaetis sibirica* (I), *Clavineta eximia* (I)

455 Parameters:

```
456 clusterenv_parser = subparsers.add_parser(name='m4cluster', parents=[parent_parser],  
457 help='Hierarchical clustering-sedimentary environment. (Module IV)\t[clustermap]')  
458 clusterenv_parser.add_argument('--level', type=str, required=True, choices=['order', 'family',  
459 'genera', 'species'], help='Taxonomic level.(For both statistical and aquatic OTUs.)')  
460 clusterenv_parser.add_argument('--aquatic', type=str2list, default=None, help='Aquatic  
461 OTUs.(default: all OTUs)\t[e.g. "OTU1,OTU2,OTU3"]')  
462 clusterenv_parser.add_argument('--geochem', type=str, required=False, help='Absolute or  
463 relative path geochemical file.(e.g. "./geochem.xlsx")')  
464 clusterenv_parser.add_argument('--output', type=str, default='./clusterenv', help='Absolute  
465 path or relative path and filename.(default: %(default)s)')  
466 clusterenv_parser.set_defaults(func=clusterenv)
```

467

468 **Command line:**

```
469 python ./TaphonomeAnalyst2.py m4cluster --input "./Supplementary material2.xlsx" --aquatic  
470 "Daohugouectes primitinus(I),Triglypta haifanggouensis,Triglypta  
471 haifanggouensis,Yanliaocorixa chinensis,Karataviella popovi,Samarura gigantea(I),Anisoptera  
472 fam. gen. sp1.(I),Platyperla platypoda(I),Ferganoconcha sibirica,Qiyia jurassica(I),Mesomyzon  
473 sp1.,Triops sp1.,Chirocephalidae gen. sp1.,Eurythoracalis mirabilis(I),Shantous  
474 lacustris(I),Foliomimus latus(I),Furvoneta viriosus(I),Furvoneta raukus(I),Mesobaetis  
475 sibirica(I),Clavineta eximia(I)" --level species
```

476

477 This module also provides Venn maps that compare biodiversity in different sedimentary
478 environments and outcrops. Venn diagrams that show differences in biodiversity across
479 various sedimentary environments or outcrops. Users need to define the taxonomic levels and
480 plot groupings as input to the Venn diagram.

481 **Parameters:**

```
482 divvenn_parser = subparsers.add_parser(name='m4venn', parents=[parent_parser],  
483 help='Venn diagram-sampling locations or environments. (Module IV)\t[venn]')  
484 divvenn_parser.add_argument('--level', type=str, default='family', choices=['order', 'family',  
485 'genera', 'species'], help='Taxonomic level.(default: %(default)s)')  
486 divvenn_parser.add_argument('--groups', type=str2list, required=True, help='Custom  
487 Groups.(Recommend to group the plots by environments or locations)\t[e.g.  
488 "plotA1/plotB2,plotB1/plotA2,plotC1/plotC2"]')  
489 divvenn_parser.add_argument('--output', type=str, default='./divvenn', help='Absolute path or
```

```

490 relative path and filename.(default: %(default)s)')
491 divvenn_parser.set_defaults(func=divvenn)
492
493 Command line:
494 python ./TaphonomeAnalyst2.py m4venn --input "./Supplementary material2.xlsx" --groups
495 Donggou1/Donggou2/Donggou3,Donggouli1/Donggouli2/Donggouli3/Beigou1/Beigou2/Beigo
496 u3 --level family

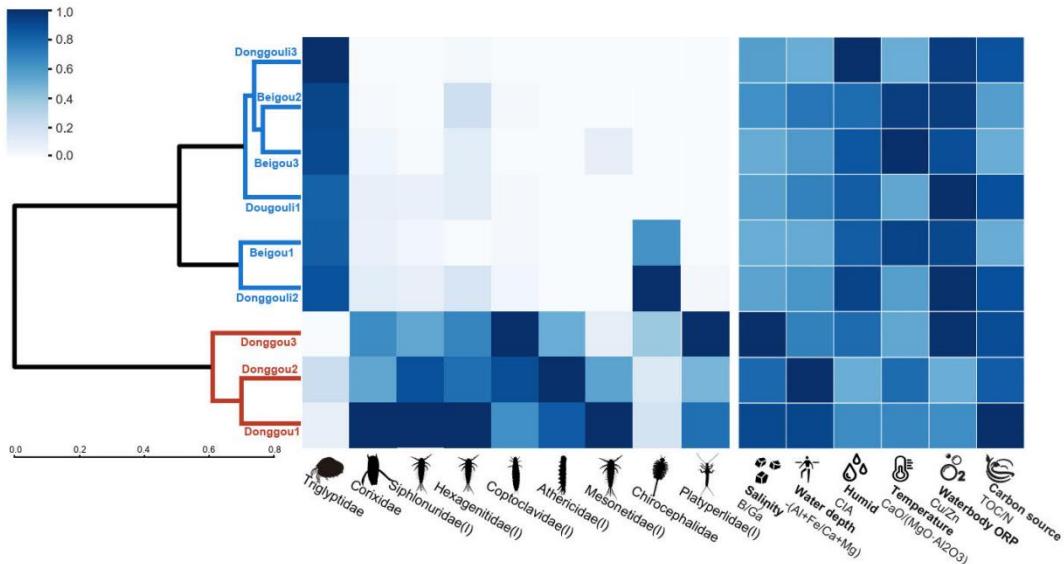
```

497 **Visualization of geochemical data (Module V)**

498 The environmental variables that shaped ancient ecosystems are not directly measurable and
499 often necessitate the extrapolation of geochemical data for accurate interpretation. The ratios
500 of elements and oxides have been widely used in ancient environmental studies to gain
501 insights into salinity, temperature, water depth, humidity, intensity of volcanic activities, and
502 other factors (Bai et al., 2020; Chen, Wan, 1999; Chen et al., 1999; Feng et al., 2003; Hou et
503 al., 2018; Hou et al., 2023; Fu et al., 2018; Harnois, 1988; Nesbitt, Young, 1982; McLennan,
504 1993; Stanistreet et al., 2020; Swain et al., 2022; Wang et al., 2023; Yang et al., 2022; Zhou
505 and Sun, 2023). Beyond serving as final resting grounds for fossilized remains, ancient water
506 bodies act as primary archives for geochemical information, frequently encapsulating
507 environmental contexts crucial to the majority of terrestrial organisms inhabiting aquatic
508 habitats to littoral zones. This module provides a geochemical heat map of different sampling
509 plots. This figure can be automatically combined with the sedimentary environment cluster
510 tree and the heat map distribution of aquatic organisms. It can simultaneously reflect the
511 distribution of biological and geochemical factors in different environmental groups.

	A	B	C	D	E	F	G
1	No.	B/Ga	CaO/(MgO -(Al+Fe)/C)	TOC/N	CIA	Cu/Zn	
2	Beigou1	1.72	0.27	(17.30)	1.67	52.17	0.88
3	Beigou2	2.22	0.27	(13.77)	2.80	51.22	0.91
4	Beigou3	1.75	0.30	(16.05)	1.60	52.59	0.86
5	Donggou1	3.32	0.17	(10.79)	9.80	49.32	0.65
6	Donggou2	2.85	0.21	(9.35)	7.00	47.06	0.53
7	Donggou3	3.70	0.13	(14.42)	8.00	51.37	0.94
8	Donggouli1	1.98	0.13	(14.45)	7.80	52.20	0.96
9	Donggouli2	1.90	0.14	(15.79)	7.80	53.77	0.95
10	Donggouli3	1.99	0.11	(17.50)	7.60	54.95	0.91

512 **Style of geochemical tables.** See Supplementary material 3 for details



513

514 **Integrative visualization based on aquatic OTU abundance and geochemical data.** This
 515 module can be automatically combined with the sedimentary environment cluster tree and the
 516 distribution heat map of aquatic organisms. It can simultaneously reflect the distribution of
 517 biological and geochemical factors in different environmental groups.

518 Parameters:

```
519 m5clusterenv_parser = subparsers.add_parser(name='m5cluster', parents=[parent_parser],  

520   help='visualization of geochemical data. (Module V)\t[clustermap]')  

521 m5clusterenv_parser.add_argument('--level', type=str, required=True, choices=['order',  

522   'family', 'genera', 'species'], help='Taxonomic level.(For both statistical and aquatic OTUs.)')  

523 m5clusterenv_parser.add_argument('--aquatic', type=str2list, default=None, help='Aquatic  

524 OTUs.(default: all OTUs)\t[e.g. "OTU1,OTU2,OTU3"]')  

525 m5clusterenv_parser.add_argument('--geochem', type=str, required=False, help='Absolute or  

526 relative path geochemical file.(e.g. "./geochem.xlsx")')  

527 m5clusterenv_parser.add_argument('--output', type=str, default='./clusterenv', help='Absolute  

528 path or relative path and filename.(default: %(default)s)')  

529 m5clusterenv_parser.set_defaults(func=clusterenv)
```

530

531 **Command line:**

```
532 python ./TaphonomAnalyst2.py m5cluster --input "./Supplementary material2.xlsx" --aquatic  

533 "Daohugouectes primitinus(l),Tritypta haifanggouensis,Tritypta  

534 haifanggouensis,Yanliaocorixa chinensis,Karataviella popovi,Samarura gigantea(l),Anisoptera  

535 fam. gen. sp1.(l),Platyperla platypoda(l),Ferganoconcha sibirica,Qiyia jurassica(l),Mesomyzon  

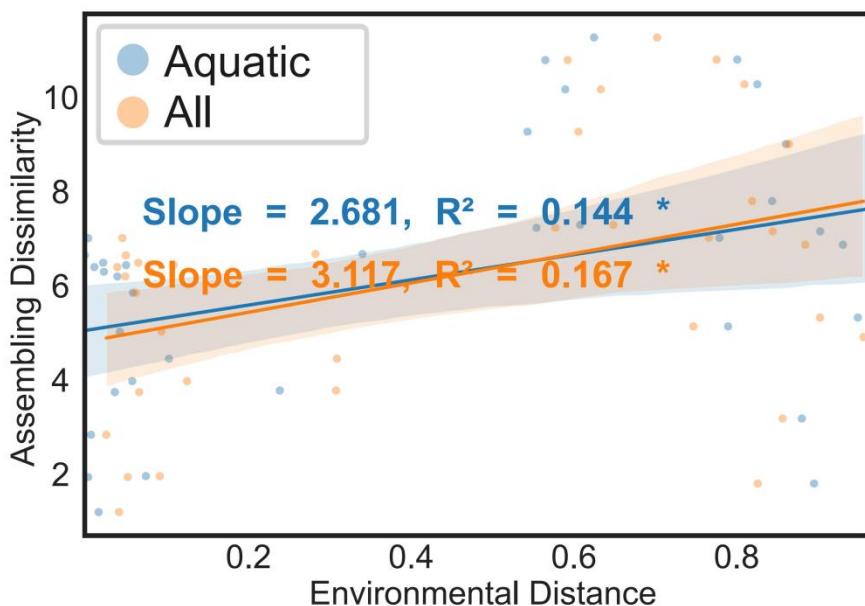
536 sp1.,Triops sp1.,Chirocephalidae gen. sp1.,Eurythoracalis mirabilis(l),Shantous  

537 lacustris(l),Foliomimus latus(l),Furvoneta viriosus(l),Furvoneta raukus(l),Mesobaetis  

538 sibirica(l),Clavineta eximia(l)" --geochem "./Supplementary material3.xlsx" --level species
```

539 **Assembling dissimilarity – environmental distance**
540 **test (Module VI)**

541 Assemblage similarities were quantified using the Bray-Curtis distance metric derived from
542 various sampling plots, whereas environmental distance was determined by employing an
543 Euclidean distance matrix based on measured geochemical variables. This module quantifies
544 the responses of aquatic and terrestrial community components to changes in environmental
545 factors. In this graph, the greater the slope of pronounced differentiation among the biological
546 abundance, the more statistically pronounced is the differentiation among the biological
547 assemblages in response to changes in environmental factors.



548
549 **Assembling dissimilarity – environmental distance test.** The slope of the line indicates the
550 degree of abrupt environmental changes corresponding to the assemblages. The relationship
551 is between faunal assembling dissimilarity and environmental distance. The statistical
552 significance is denoted by asterisks, whereby *** indicates $p < 0.001$, ** indicates $p < 0.01$, and
553 * indicates $p < 0.05$. The similarity in trends between aquatic assemblages and all assemblages
554 suggests that both are consistently responsive to environmental changes.

555
556 A list of aquatic OTUs is required for this step. The list aquatic OTUs taxonomic level must
557 correspond your research.

558 **Parameters:**

559 `dissenvtest_parser = subparsers.add_parser(name='m6dissenvtest', parents=[parent_parser],`
560 `help='Assembling dissimilarity- environmental distance test. (Module VI)\t[regplot]')`
561 `dissenvtest_parser.add_argument('--geochem', type=str, required=True, help='Absolute or`
562 `relative path geochemical file.(e.g. "./geochem.xlsx")')`
563 `dissenvtest_parser.add_argument('--aquatic', type=str2list, required=True, help='Aquatic`
564 `OTUs.\t[e.g. "OTU1,OTU2,OTU3"]')`
565 `dissenvtest_parser.add_argument('--level_aquatic', type=str, required=True, choices=['order',`

```

566 'family', 'genera', 'species'], help="Taxonomic level for aquatic OTUs.")
567 dissenvtest_parser.add_argument('--level_terrestrial', type=str, required=True,
568 choices=['order', 'family', 'genera', 'species'], help="Taxonomic level for terrestrial OTUs.")
569 dissenvtest_parser.add_argument('--output', type=str, default='./dissenvtest', help='Absolute
570 path or relative path and filename.(default: %(default)s)')
571 dissenvtest_parser.set_defaults(func=dissenvtest)
572
573 Command line:
574 python ./TaphonomAnalyst2.py m6dissenvtest --input "./Supplementary material2.xlsx" --
575 aquatic "Daohugouneutes primitinus(l),Triglypta haifanggouensis,Triglypta
576 haifanggouensis,Yanliaocorixa chinensis,Karataviella popovi,Samarura gigantea(l),Anisoptera
577 fam. gen. sp1.(l),Platyperla platypoda(l),Ferganoconcha sibirica,Qiyia jurassica(l),Mesomyzon
578 sp1.,Triops sp1.,Chirocephalidae gen. sp1.,Eurythoracalis mirabilis(l),Shantous
579 lacustris(l),Foliomimus latus(l),Furvoneta viriosus(l),Furvoneta raukus(l),Mesobaetis
580 sibirica(l),Clavineta eximia(l)" --level_aquatic species --level_terrrestrial family --geochem
581 "./Supplementary material3.xlsx"

```

582 Mantel Test between species abundance and 583 ecological environmental variables (Module VII)

584 In 1967, Nathan Mantel revolutionized statistical analysis by proposing the Mantel Test. This
585 method advanced statistics beyond the constraints of traditional correlation coefficients, which
586 then were only equipped to analyze pairwise relationships among variables within a single
587 data matrix. The Mantel Test broke new ground by facilitating the assessment of correlations
588 between two distinct matrices (Mantel,1967).

589 Since its inception, the Mantel Test has been integral to understanding the rapid evolution and
590 application across diverse scientific domains, notably in microbial community ecology. In the
591 realm of paleoecology, the Mantel Test serves as an invaluable tool for probing the
592 connections between geochemical factors and fluctuations in biological abundance. To
593 quantify similarities in species assemblages, the Bray-Curtis dissimilarity metric is commonly
594 employed and is derived from comparative data gathered from various sampling plots.
595 Concurrently, the environmental distances are defined using a Euclidean distance matrix
596 predicated on quantified geochemical variables.

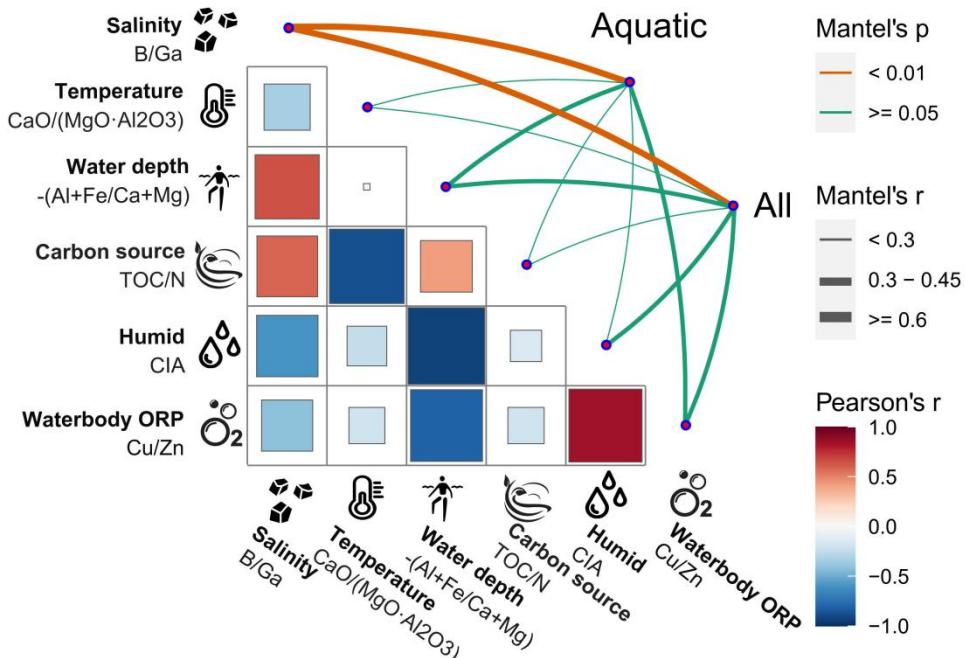
597 The Bray-Curtis distance is:

$$598 d_{(u, v)} = \frac{\sum_i |u_i - v_i|}{\sum_i |u_i + v_i|} \quad (4)$$

599

600 and the Euclidean distance is:

$$601 d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (5)$$



602

603 **Mantel test between environmental factors and palaeocommunity composition (family**
 604 **level).** The strength of the correlation is represented by the partial Mantel's r statistic, with line
 605 width indicating the magnitude of the correlation and line color denoting statistical
 606 significance. Pairwise comparisons of environmental factors were also conducted, with a color
 607 gradient representing the strength of the Pearson correlation. The Mantel test indicates that
 608 salinity is primarily responsible for significant relationships in aquatic and other assemblages
 609 across varying environmental distances.

610 Here, it is necessary to set the taxonomic level used for quantifying assemblage differences.
 611 A list of aquatic OTUs is required for this step. The list of aquatic OTUs and their taxonomic
 612 level must correspond to your research question.

613 **Parameters:**

```
614 mantel_parser = subparsers.add_parser(name='m7mantel', parents=[parent_parser],  

615 help='Mantel Test between species abundance and ecological environmental variables.  

616 (Module VII)\t[multiplot]')  

617 mantel_parser.add_argument('--rhome', type=str, required=True, help='Absolute path of  

618 R_HOME.(e.g. "C:\Program Files\R\R-4.1.3")')  

619 mantel_parser.add_argument('--geochem', type=str, required=True, help='Absolute or relative  

620 path geochemical file.(e.g. "./geochem.xlsx")')  

621 mantel_parser.add_argument('--aquatic', type=str2list, required=True, help='Aquatic  

622 OTUs.\t[e.g. "OTU1,OTU2,OTU3"]')  

623 mantel_parser.add_argument('--level_aquatic', type=str, required=True, choices=['order',  

624 'family', 'genera', 'species'], help="Taxonomic level for aquatic OTUs.")  

625 mantel_parser.add_argument('--level_terrestrial', type=str, required=True, choices=['order',  

626 'family', 'genera', 'species'], help="Taxonomic level for terrestrial OTUs.")  

627 mantel_parser.add_argument('--corr', type=str, default='pearson', choices=['pearson',  

628 'spearman', 'kendall'], help='Correlation algorithm for geochem.(default: %(default)s)')
```

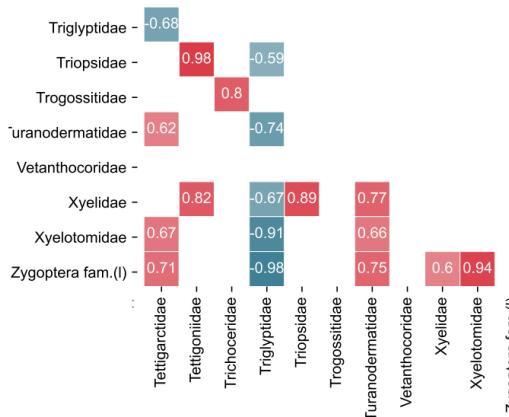
```
629 mantel_parser.add_argument('--output', type=str, default='./mantel', help='Absolute path or  
630 relative path and filename.(default: %(default)s)'  
631 mantel_parser.set_defaults(func=mantel)  
632
```

633 Command line:

```
634 python ./TaphonomeAnalyst2.py m7mantel --input "./Supplementary material2.xlsx" --rhome  
635 "C:\Program Files\R\R-4.1.3" --geochem "./Supplementary material3.xlsx" --aquatic  
636 "Daohugouectes primitinus(I),Triglypta haifanggouensis,Triglypta  
637 haifanggouensis,Yanliaocorixa chinensis,Karataviella popovi,Samarura gigantea(I),Anisoptera  
638 fam. gen. sp1.(I),Platyperla platypoda(I),Ferganoconcha sibirica,Qiyia jurassica(I),Mesomyzon  
639 sp1.,Triops sp1.,Chirocephalidae gen. sp1.,Eurythoracalis mirabilis(I),Shantous  
640 lacustris(I),Foliomimus latus(I),Furvoneta viriosus(I),Furvoneta raukus(I),Mesobaetis  
641 sibirica(I),Clavineta eximia(I)" --level_aquatic species --level_terrestrial family --corr pearson
```

642 Species correlation semi-matrix graphics (Module VIII)

```
643 The module for species correlation analysis aims to uncover possible interactions among  
644 fossil species, which are indicative of symbiotic relationships. TaphonomeAnalyst 2.0 offers a  
645 variety of techniques to compute correlations and to generate semi-matrix graphics, including  
646 Pearson's, Spearman's, Kendall's rank, and SparCC correlation coefficients.
```



647 Semi-matrix of taphonomy correlations among organisms (a part of an entire figure).

```
648 Red indicates a positive correlation, blue a negative correlation. Missing cells are due to  
649 filtering of data with insufficient significance levels. Users can adjust the intensity of data  
650 filtering as needed.
```

```
651  
652 Here the user is required to enter the taxonomic rank, correlation type and intensity, and P-  
653 value.
```

```
654 Parameters:
```

```
655 corrotus_parser = subparsers.add_parser(name='m8corrotus', parents=[parent_parser],  
656 help='Heatmap-OTUs correlation analysis. (Module VIII)\t[heatmap]')  
657 corrotus_parser.add_argument('--level', type=str, default='family', choices=['order', 'family',  
658 'genera', 'species'], help='Taxonomic level.(default: %(default)s)')  
659 corrotus_parser.add_argument('--corr', type=str, default='pearson', choices=['pearson',
```

```
660 'spearman', 'kendall', 'sparcc'], help='Correlation algorithm.(default: %(default)s)')  
661 corrotus_parser.add_argument('--p_value', type=float, default=0.1, help='Maximum threshold  
662 of p-value.(default: %(default)s)')  
663 corrotus_parser.add_argument('--output', type=str, default='./corrotus', help='Absolute path or  
664 relative path and filename.(default: %(default)s)')  
665 corrotus_parser.set_defaults(func=corrotus)  
666  
667 Command line:  
668 python ./TaphonomicAnalyst2.py m8corrotus --input "./Supplementary material2.xlsx" --level  
669 family --corr pearson --p_value 0.1
```

670 Correlational Network Visualization (Module IX)

671 IX, 1 Overview

672 A network characterized by nodes and links is a composition of both elements. In such a
673 network, the nodes symbolize taxa found within sampling plots, while the edges represent the
674 taphonomic co-occurrences among these taxa. Animals inhabiting the same
675 microenvironment also demonstrate co-occurrence during the process of fossilization.
676 Therefore, the phenomenon of taphonomic co-occurrence partly can be explained by the
677 extent of habitat overlap between two taxa. In constructing the network, two crucial
678 considerations must be addressed: the methodology for correlation calculation and division of
679 network modules.

680 IX.2 Correlation

681 Module links include various correlation methods, such as Pearson's correlation, Spearman's
682 rank correlation, Kendall's rank correlation, and SparCC (sparse correlations for
683 compositional data) coefficients for network visualization. It also enables users to define their
684 own filters, including correlation strength and P-values.
685 The most widely employed method in microbial research is SparCC correlation coefficient,
686 which is a computational method designed to identify correlations between specifications
687 within a community by analyzing sequential data (Kurtz et. al., 2015). SparCC does not use
688 variance as a direct measure of correlation, but instead adopts a more complex method to
689 estimate the correlation between species with sparse data. SparCC improves the estimation
690 of correlation between microbial abundances by log-transforming the observed abundance
691 data and using the variance of log ratios to correct for biases. The method also employs a
692 bootstrap procedure to calculate P-values, allowing for assessment of the statistical
693 significance of the correlations. However, its performance is constrained when computing
694 interaction networks for high dimensional datasets. Typically, microbial network studies
695 require a minimum of 3000 samples within the total dataset, with a OTUs diversity not
696 exceeding 800.
697 SparCC is based on the log-ratio transformation:

698 $y_{ij} = \log \frac{x_i}{x_j} = \log x_i - \log x_j$ (6)

699 x_i : the fraction of OTU i. x_j : the fraction of OTU j.

700 Aitchison proposed using the quantity where the variance is taken across all samples to
701 describe the dependencies in a compositional dataset (Kurtz et. al., 2015).

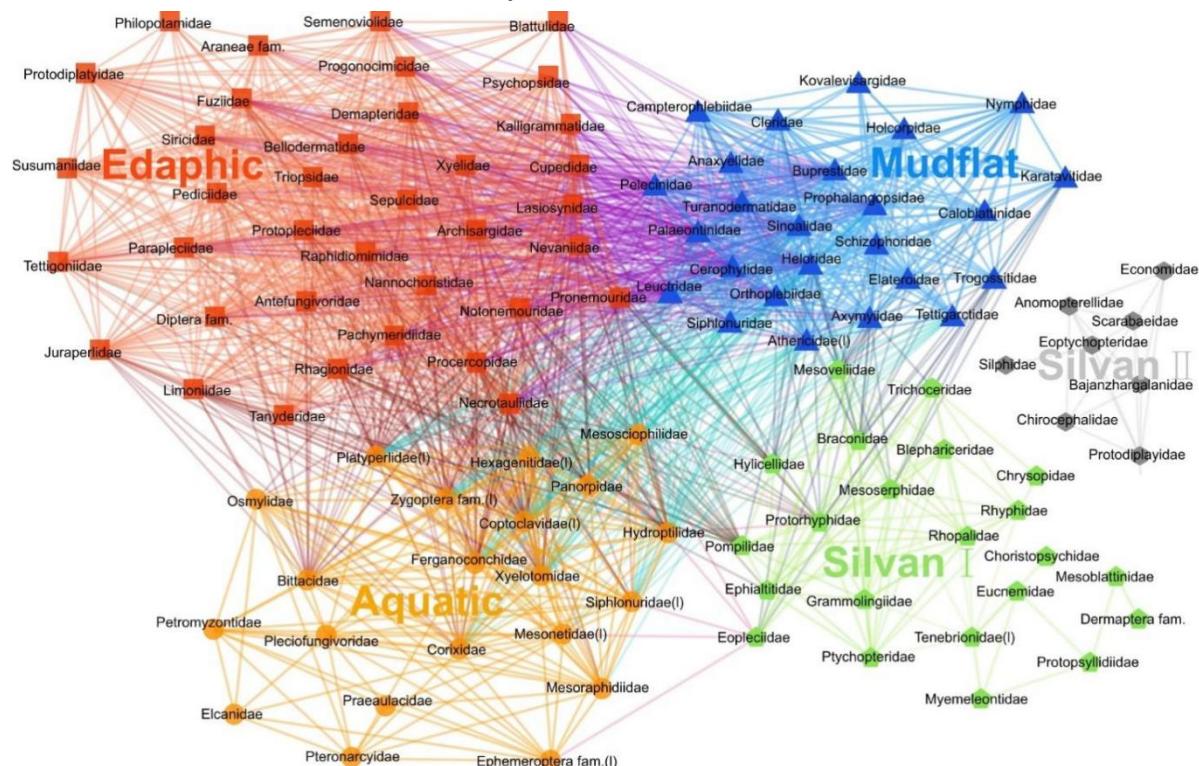
702 $t_{ij} = \text{Var}[\frac{x_i}{x_j}] = \text{Var}[y_{ij}]$ (7)

703 When OUT i and j are absolutely correlated and their ratio is constant; consequently, $t_{ij}=0$

704 and $t_{ij} = w_i^2 + w_j^2 - 2\rho_{ij}w_iw_j$

705 w_i^2 , w_j^2 : The variances of the log-transformed basis abundances OUT i and j. ρ_{ij} : the

706 correlation between them OUT i and j.



707 **Network visualization.** Points of the same color represent the same module, indicating that
708 these organisms likely inhabit the same environment. The output image is in PDF format,
709 allowing users to adjust the font size in a PDF editor to meet publishing requirements. Based
710 on randomization tests, this threefold community structure is statistically significant (Pearson
711 correlation > 0, P < 0.1). This algorithm segregates families into five modules, which also
712 means five environments. The first is an aquatic assemblage (orange) that includes water
713 boatman (Corixidae); mayfly naiads of Mesonetidae, Siphlonuridae, and Hexagenitidae,
714 naiads of stoneflies and dragonflies, *Daohugouneutes primitinus*, *Ferganoconcha sibirica* and
715 a lamprey genus in the Petromyzontidae. This assemblage represents a deep-water
716 environment. Second, is a mudflat assemblage (blue), that includes adults of soldier flies

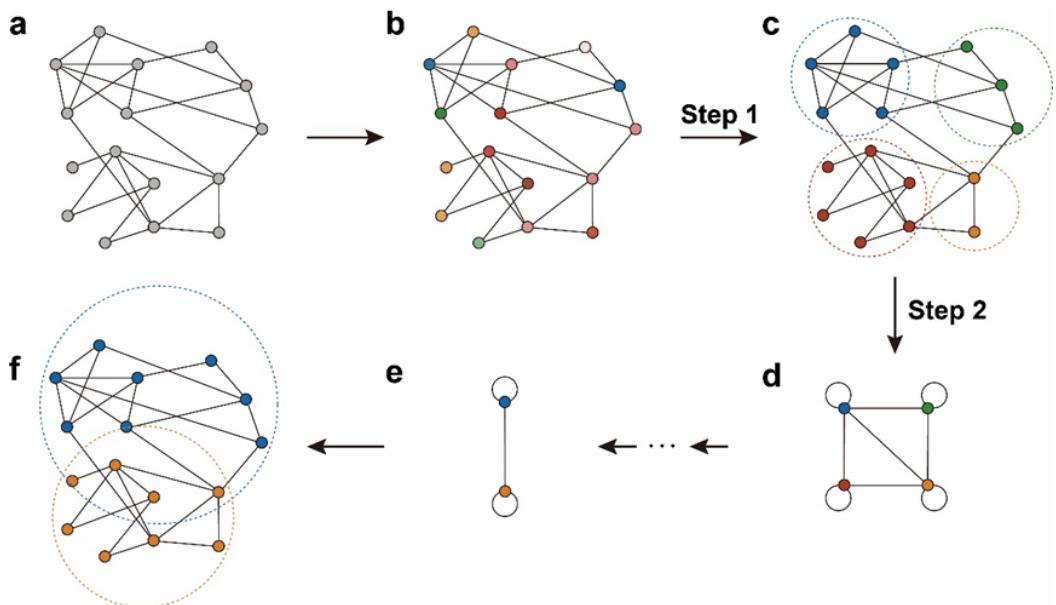
717 (Kovalevisargidae), axymyiids (Axymyiidae), snipe flies (Rhagionidae), primitive crane flies
718 (Tanyderidae), sinoalids (Sinoalidae); and adults of aquatic insects such as primitive minnow
719 mayflies (Siphlonuridae) and rolled-winged stoneflies (Leuctridae). This assemblage is typical
720 of wet environments colonized by herbaceous plants adjacent to water bodies. Third, is the
721 edaphic assemblage (red) that consists generally of cockroaches of the Fuziidae and
722 Blattulidae; earwigs of the Bellodermatidae, Dermapteridae, and Turanodermatidae; and a
723 beetle assigned to Lasiosynidae. This assemblage expressed an affinity for an above-ground
724 environment of humus and litter of ambient vegetation. Fourth, is the silvan assemblage
725 (bright green) that includes a green lacewing (Chrysopidae), a wood gnat (Rhyphidae), an
726 antlion (Myrmeleontidae), a lacewing (Grammalingiidae), and a primitive, cicada-like form
727 (Protopsylidiidae). This assemblage shows a preference for a relatively dry environment. A
728 small assemblage (grey), including scarab beetles (Scarabaeidae), caddisflies (Economidae),
729 and primitive crane flies (Eoptychopteridae), probably inhabited forests.

730 **IX. 3 Louvain algorithm**

731 The Louvain algorithm identifies communities based on the concept of modularity. When there
732 is a high density of connections within a module and a low density of connections between
733 modules, the network exhibits high modularity as a result of this partitioning. The iteration
734 process ceases automatically when there is no additional increase in modularity. Modularity is
735 the set $Q \subseteq [-0.5, 1]$, and it can also be used to evaluate the effectiveness of network module
736 division. The larger the value, the better the module division effect. When the modularity is
737 between 0.3 and 0.7, it indicates that the clustering effect is very good. The formula for
738 modularity is:

739
$$Q = \frac{1}{2m} \sum_{v \omega} \left[A_{v \omega} - \frac{k_v k_\omega}{2m} \right] \delta(c_v, c_\omega) \quad (8)$$

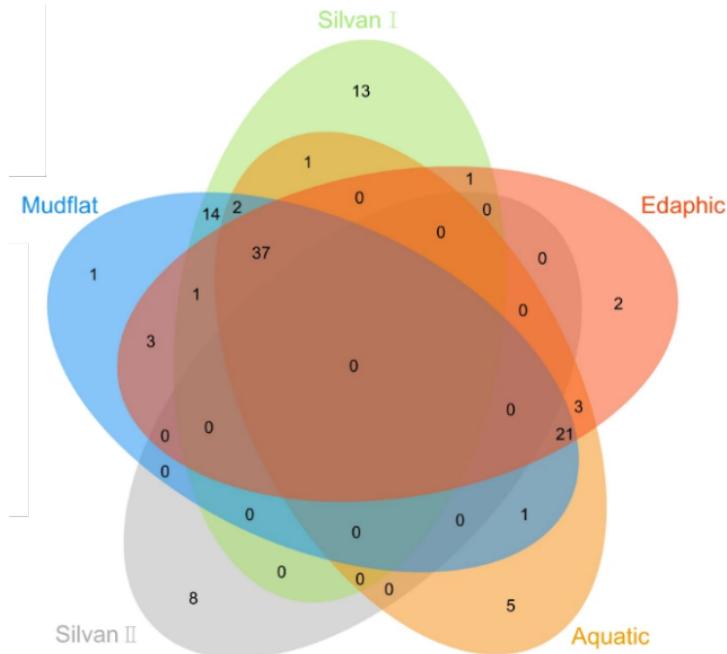
740 m indicates the total number of connections of all nodes. v and ω are two nodes in the
741 network. $A_{v \omega} = 0$, if v and ω are not linked. $A_{v \omega} = 1$, if v and ω are linked. k_v = degree of the
742 node v . k_ω = degree of the node ω . $\delta(c_v, c_\omega) = 0$, if v and ω are not linked. $\delta(c_v, c_\omega) = 1$, if v and
743 ω are linked.



744 **The operation procedure of Louvain algorithm.** **a**, The network is constructed based on
 745 correlation matrix. **b, c**, Step one: each node in the network is assigned to its individual
 746 module. The Louvain algorithm examines the increase in modularity if neighboring nodes of
 747 node u are assigned to the same module. If there is no increase in modularity results when
 748 neighboring nodes joined (i.e., the gain in modularity is zero or negative), then node-
 749 neighboring nodes remain in their original module. This procedure is repeatedly carried out
 750 until no further enhancement in modularity can be achieved by altering the community
 751 assignments of any nodes, at which point the first phase is concluded. **c, d**, Step two: after
 752 the algorithm divides the modules in the first round, each module is merged into a large self-
 753 looping node. The weight of the self-looping node is the sum of the link weights of all nodes in
 754 the original module. The newly merged nodes's weight is the sum of the link weights of all
 755 nodes in the original module. **d, e**, The subsequent calculation method is the same as step
 756 one. **f**, The final output of the resulting graph. The increase in community modularity can be
 757 computed using the following formula:

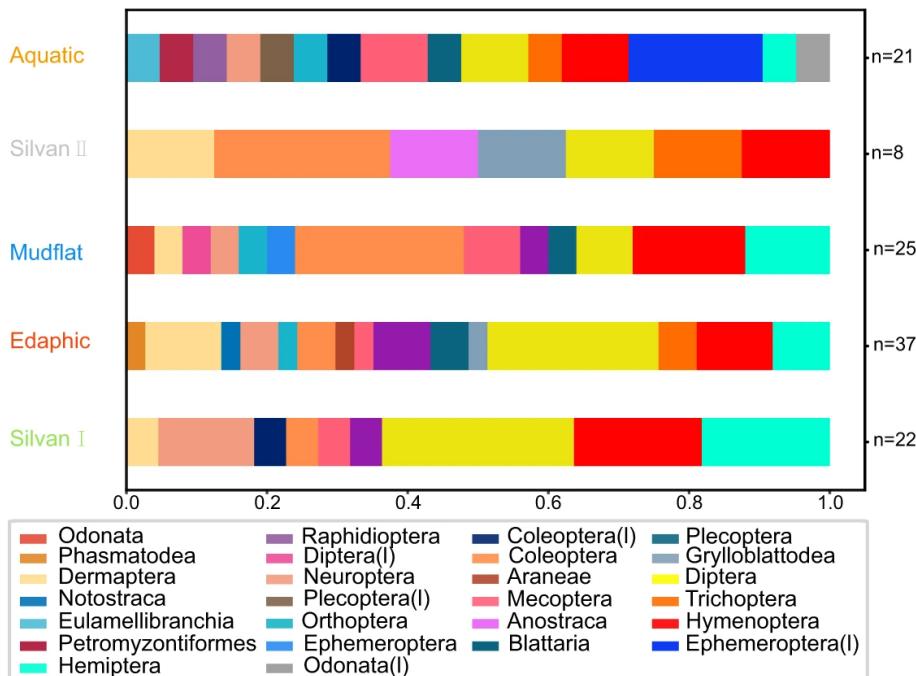
758
$$\Delta Q = \left[\frac{\sum_{\text{in}} + k_{v,\text{in}}}{2m} - \left(\frac{\sum_{\text{tot}} + k_v}{2m} \right)_2 \right] - \left[\frac{\sum_{\text{in}}}{2m} - \left(\frac{\sum_{\text{tot}}}{2m} \right)_2 - \left(\frac{k_v}{2m} \right)^2 \right] \quad (9)$$

759 \sum_{in} is the sum of the weights of all links in same module. \sum_{tot} is the sum of all
760 links that are external to the module. k_v is the sum of weights of node v . $k_{v,\text{in}}$ is the sum of



761 weights between node v to the module whereby node v is moved.

762 **Venn map of biodiversity between different modules is output with the network.** The
763 diversity and intersections of different modules in the network can be viewed.



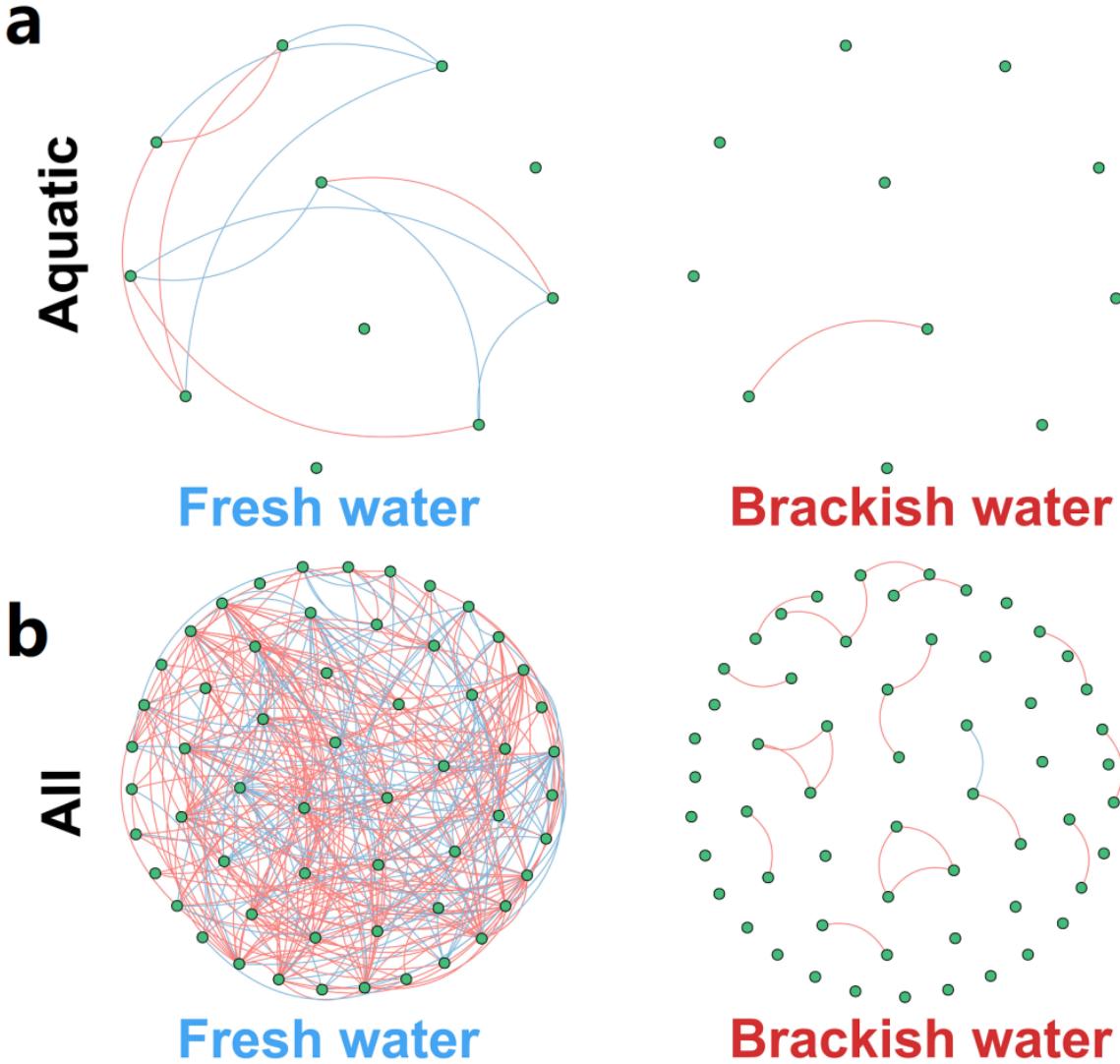
764

765 **A histogram of the biodiversity for each module is output with the network.**

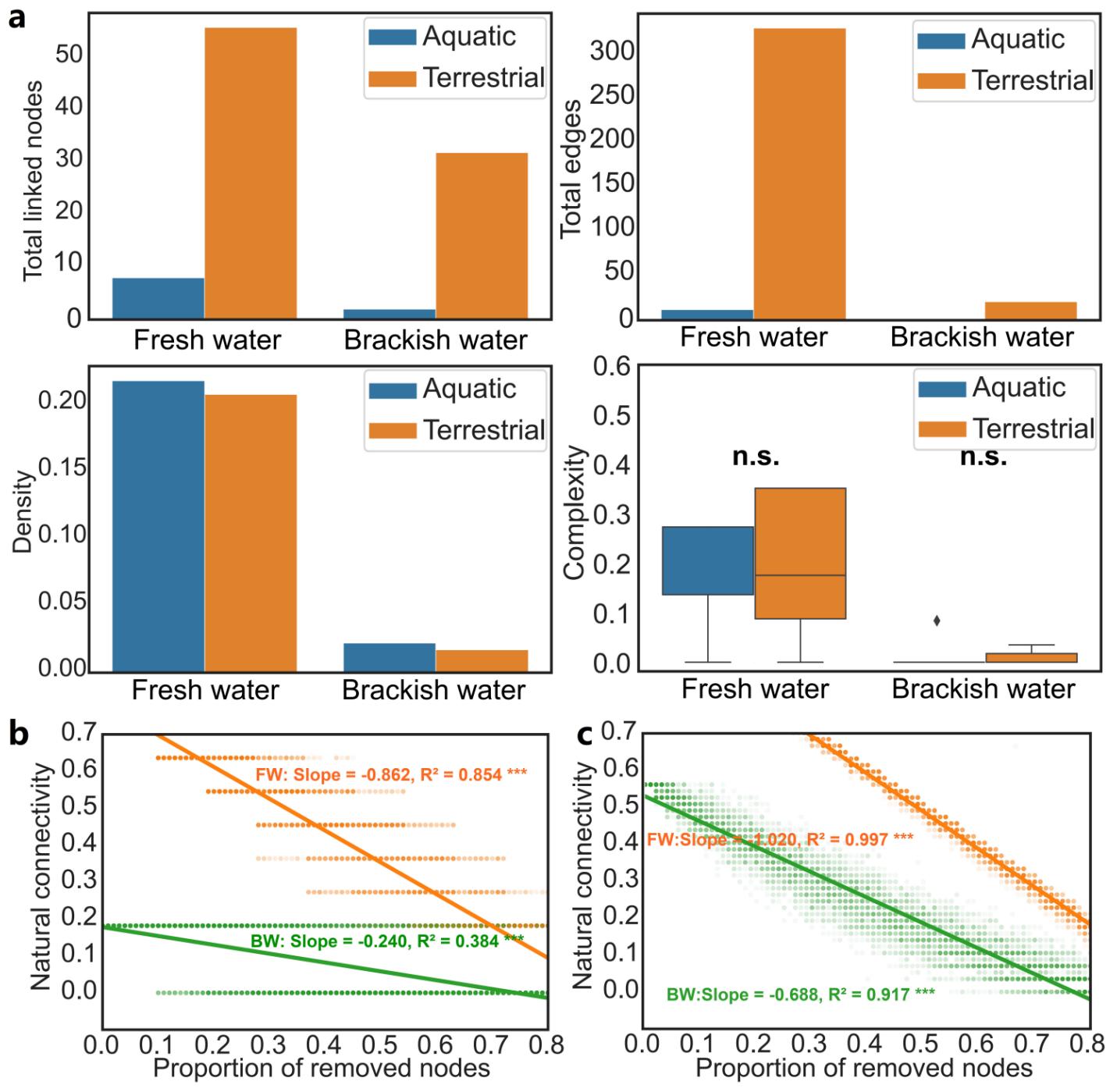
766 Users need to input the taxonomic level, types and strength conditions of the correlations, and
767 p-value conditions.
768 Parameters:
769 cooccurnet_parser = subparsers.add_parser(name='m9cooccurnet', parents=[parent_parser],
770 help='Co-occurrence networks. (Module IX)\n[network/venn]')
771 cooccurnet_parser.add_argument('--level', type=str, default='family', choices=['order', 'family',
772 'genera', 'species'], help='Taxonomic level.(default: %(default)s)')
773 cooccurnet_parser.add_argument('--corr', type=str, default='pearson', choices=['pearson',
774 'spearman', 'kendall', 'sparcc'], help='Correlation algorithm.(default: %(default)s)')
775 cooccurnet_parser.add_argument('--corr_coef', type=float, default=0.7, help='Minimum
776 threshold of correlation coefficient.(default: %(default)s)')
777 cooccurnet_parser.add_argument('--p_value', type=float, default=0.1, help='Maximum
778 threshold of p-value.(default: %(default)s)')
779 cooccurnet_parser.add_argument('--output', type=str, default='./cooccurnet', help='Absolute
780 path or relative path and filename.(default: %(default)s)')
781 cooccurnet_parser.set_defaults(func=cooccurnet)
782
783 Command line:
784 python ./TaphonomAnalyst2.py m9cooccurnet --input "./Supplementary material2.xlsx" --level
785 family --corr pearson --corr_coef 0.7 --p_value 0.1

786 Comparison of networks (Module X)

787 This module offers the capability to compare networks under different groups of sampling
788 plots. The network set for comparison can be easy to read by using a consistent, random
789 layout. Additionally, it presents a bar chart that contrasts key network metrics such as total
790 nodes, total linked nodes, total edges, density, modularity, complexity, degree, and
791 robustness.
792 The network set for comparison can be easy to read by using a consistent, random layout. In
793 this plot, points at the same location within the network represent the same species. The
794 network enables users to discern differences in community correlations by examining the
795 number of linked points and the connection density of links. The grouping is based on the
796 cluster analysis above. The increase of nutrient salinity decreases the complexity and
797 modularity of the network.



798 **Co-occurrence patterns in the Daohugou fauna across contrasting salinity gradients.**
 799 The diversity of aquatic assemblages is calculated at the species level. Nodes represent
 800 individual operational taxonomic units, with their color and size positively correlated to the
 801 nodes. The edges signify significant Spearman correlations (correlation > 0.7 , $p < 0.01$), with
 802 red lines indicating positive correlations and blue lines representing negative correlations. **a**,
 803 Co-occurrence network of the aquatic community. **b**, Co-occurrence network of the all
 804 communities.



805 **Network metrics and robustness evaluations.** Significance level is sn.s. >0.05, *p<0.05,
 806 **p<0.01 and ***p<0.001. **a**, Statistical indicators of networks under different salinities. **b**, The
 807 robustness for the aquatic network. **c**, The robustness for the all network.

If you encounter any issues, please feel free to reach out to Wang Ma(马旺) (wma19952022@163.com).

808 (1) First, input the sample grouping; if not the set; the default is the sedimentary environment
809 clustering result.

810 (2) Second, a list of aquatic OTUs is required for this step. The list aquatic OTUs and their
811 taxonomic rank must correspond your research.

812 (3) Third, users need to input the taxonomic rank, types and strength of conditions of
813 correlations, and p-value conditions.

814 **Parameters:**

```
815 netVC_parser = subparsers.add_parser(name='m10netVC', parents=[parent_parser],  
816 help='Generate a unified layout network for comparison. (Module  
817 X)\t[network/boxplot/barplot]')  
818 netVC_parser.add_argument('--aquatic', type=str2list, required=True, help='Aquatic  
819 OTUs.\t[e.g. "OTU1,OTU2,OTU3"]')  
820 netVC_parser.add_argument('--level_aquatic', type=str, required=True, choices=['order',  
821 'family', 'genera', 'species'], help="Taxonomic level for aquatic OTUs.")  
822 netVC_parser.add_argument('--level_terrestrial', type=str, required=True, choices=['order',  
823 'family', 'genera', 'species'], help="Taxonomic level for terrestrial OTUs.")  
824 netVC_parser.add_argument('--groups', type=str2list, required=True, help='Environment  
825 groups.(Grouping the plots by different aquatic and terrestrial environments)\t[e.g.  
826 "plotA1/plotB2,plotB1/plotA2,plotC1/plotC2"]')  
827 netVC_parser.add_argument('--corr', type=str, default='pearson', choices=['pearson',  
828 'spearman', 'kendall', 'sparcc'], help='Correlation algorithm.(default: %(default)s)')  
829 netVC_parser.add_argument('--corr_coef', type=float, default=0.7, help='Minimum threshold  
830 of correlation coefficient.(default: %(default)s)')  
831 netVC_parser.add_argument('--p_value', type=float, default=0.1, help='Maximum threshold of  
832 p-value.(default: %(default)s)')  
833 netVC_parser.add_argument('--output', type=str, default='./netVC', help='Absolute path or  
834 relative path and filename.(default: %(default)s)')  
835 netVC_parser.set_defaults(func=netVC)
```

836 **Command line:**

```
837 python ./TaphonomAnalyst2.py m10netVC --input "./Supplementary material2.xlsx" --aquatic  
838 "Daohugouectes primitinus(I),Triglypta haifanggouensis,Triglypta  
839 haifanggouensis,Yanliaocorixa chinensis,Karataviella popovi,Samarura gigantea(I),Anisoptera  
840 fam. gen. sp1.(I),Platyperla platypoda(I),Ferganoconcha sibirica,Qiyia jurassica(I),Mesomyzon  
841 sp1.,Triops sp1.,Chirocephalidae gen. sp1.,Eurythoracalis mirabilis(I),Shantous  
842 lacustris(I),Foliomimus latus(I),Furvoneta viriosus(I),Furvoneta raukus(I),Mesobaetis  
843 sibirica(I),Clavineta eximia(I)" --level_aquatic species --level_terrestrial family --groups  
844 Donggou1/Donggou2/Donggou3,Donggouli1/Donggouli2/Donggouli3/Beigou1/Beigou2/Beigo  
845 u3 --corr spearman --corr_coef 0.7 --p_value 0.01
```

If you encounter any issues, please feel free to reach out to Wang Ma(马旺) (wma19952022@163.com).

Evaluation indicators	Significance in netology	Significance in ecology
Total nodes	All nodes present in the network, both connected and unconnected.	The diversity present in the sampling plots.
Total linked nodes	All connected nodes present in the network.	The diversity present clear co-occurrences with other OTUs.
Total edges	All links present in the network.	Taphonomic co-occurrences. It can also reflect the symbiotic relationships among OTUs to a certain extent.
Density	The density of interconnecting edges between nodes in a network. $d(G) = \frac{2L}{N(N - 1)}$ (10) N: Total nodes. L: Total edges.	The complexity of a symbiotic relationship.
Complexity	The complexity of the network. $C_i = \frac{d_i}{n}$ (11) d _i : degree of node i. n: total nodes. C _i : complexity of node i.	Community complexity.
Degree	The centrality and importance of the network. $\deg(v) = \sum_{u \in V, u \neq v} A_{uv}$ (12) V: All the nodes in the network. A: A= 1 if there is an edge from node u to node v, otherwise A=0.	The importance of keystone species in communities
Modularity	The strength of a network divided into modules. A network with high modularity has dense connections between nodes within a module, but sparse connections between nodes among different modules.	The degree of differentiation of the community.
Robustness	Response of network links after nodes are randomly removed.	The ability of communities to resist environmental change

846

847 **Table 3. Evaluation indicators of the network.**

848 Frequently Asked Questions

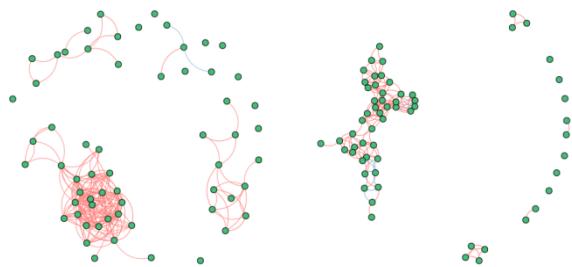
849 **The error of identical OTUs being identified as two 850 separate ones.**

851 This is often caused by incorrect spacing after or in the scientific names of the OTUs.

852 **The proportion of taphonomic grades error**

853 The primary reason for the inability to output images is that this module only accepts
854 uppercase letters A through E. Careful proofreading during input is essential, and it is
855 recommended to use Excel's statistical functions to eliminate incorrect data.

856 **The network comparison error**



857
858 The network comparison visualization may face challenges due to an overabundance of links.
859 This plethora of connections complicates the software's ability to identify matching layouts.
860 We suggest implementing the spearman correlation coefficient or imposing stricter correlation
861 criteria to diminish the link count.

862 **The geochemical heat map error**

863 In geochemical data, the concentration of certain elements may be below the detection limit,
864 which is generally written as BDL in the main text of the paper. However, in the geochemical
865 tables used for calculations, numerical values must be entered. We have referred to some
866 environmental science papers and suggest using half of the detection limit for such a value.

867 Future developments

868 TaphonomeAnalyst 2.0 is an ongoing project that will remain in a state of near-term
869 development. The software is designed to keep pace with the ever-evolving field of
870 geochemistry and taphonomy, ensuring that it remains current and relevant as the discipline
871 advances. To comprehensively grasp biological behaviors and interactions, it is essential to
872 gather, integrate, and analyze multiple types of ecological data. Future advancements in
873 TaphonomeAnalyst 2.0 will be geared towards a more holistic analyses that encompass
874 various forms of palaeoecological data, including functional morphology, herbivore arthropod
875 and pathogen damage types, sporopollen taxa, and body size. The objective of
876 TaphonomeAnalyst 3.0 would be to facilitate the creation of multilayer ecological networks
877 through the straightforward use of a diverse set of fossil community-level data.
878 The development environment of software is based on Python 3.9.21 and R 4.1.3, with the
879 following third-party libraries used in Python: matplotlib, numpy, pandas, igraph, networkx,
880 seaborn, scipy, community, skbio, h5py, numba, dask, venn, rpy2. The implementation of the
881 SparCC algorithm originates from the micnet library. In R, the third-party libraries include
882 dplyr, linkET, and ggplot2.

883 Acknowledgments

884 The original version of TaphonomeAnalyst was published in *Communications Biology*. We
885 extend our gratitude to Luke R. Grinham, the Primary Handling Editor at *Communications*
886 *Biology*, for his instrumental role in facilitating peer reviews of our original version from two
887 experts in stratigraphy and one in microbiology. Additionally, we express our sincere
888 appreciation to the three initial reviewers of the original version of TaphonomeAnalyst
889 manuscript: Roxanne Banker, Alexander Dunhill, and an anonymous reviewer for their
890 valuable insights. This work was supported by grants from the National Natural Science
891 Foundation of China (42288201 and 32020103006) and High-level Teachers in Beijing
892 Municipal Universities (no. BPHR20220114).

893 References

- 894 Allison, P.A., Briggs, D.E.G., 1991. Taphonomy of nonmineralized tissues, in *Taphonomy:*
895 *Releasing the Data Locked in the Fossil Record*. Springer. Press, pp. 25–70.
896 Bai, H., Kuang, H., Liu, Y., Peng, N., Chen, X., Wang, Y., 2020. Marinoan-aged red beds at
897 Shennongjia, South China: Evidence against global-scale glaciation during the
898 Cryogenian. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **559**, 109967.
899 Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.C.C., Charles, T., Chen, X.,
900 Cocolin, L., Eversole, K., Corral, G.H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A.,
901 Macklin, J.A., Maguin, E., Mauchline, T., McClure, R., Mitter, B., Ryan, M., Sarand, I.,
902 Smidt, H., Schelkle, B., Roume, H., Kiran, G.S., Selvin, J., Souza, R.S.C., Overbeek, L.V.,

- 903 Singh, B.K., Wagner, M., Walsh, A., Sessitsch, A., Schloter, M., 2020. Microbiome
904 definition re-visited: old concepts and new challenges. *Microbiome* **8**, 1–22.
905 <https://doi.org/10.1186/s40168-020-00875-0>
- 906 Blondel, V., Guillaume, J.L., Lambiotte, R., Lefebvre, E., 2008. Fast Unfolding of Communities
907 in Large Networks. *J. Stat. Mech.-Theory. E.* P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>.
- 908 Chao, A., 1984. Non-parametric estimation of the number of classes in a population. *Scand. J. Stat.* **11**, 265–270.
- 909 Chao, A., Lee, S.M., 1992. Estimating the number of classes via sample coverage. *J. Am. Stat. Assoc.* **87** (417), 210–217.
- 910 Chao, A., Shen, T. J., 2004. Nonparametric prediction in species sampling. *J. Agr. Biol. Envi. St.* **9**, 253–269.
- 911 Chao, A., Yang, M.C. K., 1993. Stopping rules and estimation for recapture debugging with
912 unequal failure rates. *Biometrika* **80** (1), 193–201.
- 913 Chen, J., Wan, G., 1999. Sediment particle size distribution and its environmental significance
914 in Lake Erhai, Yunnan province. *Chin. J. Geochem.* **18**(4), 314–320.
- 915 Chen, J., Wan, G., Chen, Z., Huang, R., 1999. Chemical elements in sediments of Lake Erhai
916 and palaeoclimate evolution. *Geochimica* **28**(6), 562–570.
- 917 Dhariwal, A., Chong, J., Habib, S., King, I.L., Agellon, L.B., Xia, J., 2017. Microbiomeanalyst:
918 a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome
919 data. *Nucleic. Acids. Res.* **45** (W1), W180–W188.
- 920 Feng, L., Chu, X., Zhang, Q., 2003. CIA (chemical index of alteration) and its applications in
921 the Neoproterozoic clastic rocks. *Earth Sci. Front.* **10**(4), 539–544.
- 922 Fu, J., Li, S., Xu, L., Niu, X., 2018. Paleo-sedimentary environmental restoration and its
923 significance of Chang 7 Member of Triassic Yanchang Formation in Ordos Basin, NW
924 China. *Pet. Explor. Dev.* **45**(6), 998–1008.
- 925 Fürsich, F. T., Aberhan, M., 1990. Significance of time-averaging for palaeocommunity
926 analysis. *Lethaia* **23**, 143–152.
- 927 Guo, S., Ma, W., Tang, Y., Chen, L., Wang, Y., Cui, Y., Liang, J., Li, L., Zhuang, J., Gu, J., Li,
928 M., Fang, H., Lin, X., Shih, C.K., Labandeira, C.C., Ren, D., 2023. A new method for
929 examining the co-occurrence network of fossil assemblages. *Commun. Biol.* **6**(1), 1102.
930 <https://doi.org/10.1038/s42003-023-05417-6>
- 931 Harnois, L., 1988. The C.I.W Index: a new chemical index of weathering. *Sediment. Geol.*
932 **55**(3), 319–322.
- 933 Hou, Q., Jin, Q., Niu, C., Zhang, R., Chen, F., Xu, J., Zhang, F. J., 2018. Distribution
934 characteristics and main controlling factors of main hydrocarbon sourcerocks in Liaodong
935 bay area. *Earth Sci.* **43**(6), 2160–2171.
- 936 Karr, J., Clapham, M., 2015. Taphonomic biases in the insect fossil record: Shifts in
937 articulation over geologic time. *Paleobiology* **41**(1), 16–32.
- 938 Kurtz, Z.D., Müller, C.L., Miraldi, E.R., Littman, D.R., Blaser, M.J., Bonneau, R.A., 2015
939 Sparse and Computationally Robust Inference of Microbial Ecological Networks. *PLoS Comput. Biol.* **11**(5): e1004226. <https://doi.org/10.1371/journal.pcbi.1004226>
- 940 Mantel, N. (1967). The detection of disease clustering and a generalized regression
941 approach. *Cancer Research*, **27**, 209–220.

- 947 McLennan, S.M., 1993. Weathering and global denudation. *J. Geol.* **101**(2), 295–303.
- 948 McNamara, M.E., Briggs, D.E., Orr, P.J., 2012. The controls on the preservation of structural
949 color in fossil insects. *Palaios* **27**(7), 443–454.
- 950 Muscente, A.D., Bykova, N., Boag, T.H., Buatois, L.A., Mángano, M.G., Eleish, A., Prabhu, A.,
951 Pan, F., Meyer, M.B., Schiffbauer, J.D., Fox, P., Hazen, R.M., Knoll, A.H., 2019. Ediacaran
952 biozones identified with network analysis provide evidence for pulsed extinctions of early
953 complex life. *Nat. Commun.* **10**(1), 911.
- 954 Nesbitt, H.W., Young, G.M., 1982. Early Proterozoic climates and plate motions inferred from
955 major element chemistry of lutites. *Nature* **299**, 715–717.
- 956 Newman, M. E. J., 2006. Modularity and community structure in networks. *Proc. Natl. Acad.
957 Sci.* **103**(23), 8577–8582.
- 958 Olszewski, T., 1999. Taking advantage of time-averaging. *Paleobiology*, **25**(2), 226–238.
- 959 Smith, D.M., Moe-Hoffman, A.P., 2007. Taphonomy of Diptera in Lacustrine Environments: A
960 Case Study from Florissant Fossil Beds, Colorado. *Palaios* **22**(6), 623–629.
- 961 Staff, G.M., Powell, E.N., 1998. The palaeoecological significance of diversity: the effect of
962 time averaging and differential preservation on macroinvertebrate species richness in
963 death assemblages. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **63**(1–3), 73–89.
- 964 Stanistreet, I.G., Boyle, J.F., Stollhofen, H., Deocampo, D.M., Deino, A., McHenry, L.J., Toth,
965 N., Schick, K., Njau, J.K., 2020. Palaeosalinity and palaeoclimatic geochemical proxies
966 (elements Ti, Mg, Al) vary with Milankovitch cyclicity (1.3 to 2.0 Ma), OGCP cores,
967 Palaeolake Olduvai, Tanzania. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **546**, 109656.
- 968 Swain, A., MacCracken, S.A., Fagan W.F., Labandeira, C.C., 2022. Understanding the ecology
969 of plant-insect interactions in the fossil record through bipartite networks. *Paleobiology*
970 **48**(2), 239–260.
- 971 Wang, S., Hethke, M., Wang, B., Tian, Q., Yang, Z., Jiang, B., 2019. High-resolution
972 taphonomic and palaeoecological analyses of the Jurassic Yanliao Biota of the Daohugou
973 area, northeastern China. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **530**, 200–216.
- 974 Wang, Y., Xu, Z., Hang, L., Xing, T., Hang, W., 2023. Lycoptera Fossil Elemental Imaging and
975 Paleoenvironmental Research via Laser Ionization Time-of-Flight Mass Spectrometry. *At.
976 Spectrosc.* **44**(2), 55–64.
- 977 Whipps, J.M., Lewis, K., Cooke, R.C., 1988. Mycoparasitism and plant disease control. In:
978 *Fungi in Biological Control Systems* (M. N. Burge, ed.). Manchester University Press, pp.
979 161–187.
- 980 Wing, S.L., Sues, H.D., Potts, R., DiMichele, W.A., Behrensmeyer, A.K., 1992. Evolutionary
981 paleoecology. In: *Terrestrial Ecosystems through Time: Evolutionary Paleoecology of
982 Terrestrial Plants and Animals* (A.K. Behrensmeyer, J.D. Damuth, W.A. DiMichele, R.
983 Potts, H.-D. Sues, and S.L. Wing, eds.) The University of Chicago Press, pp. 1–13.
- 984 Wright, P., Cherns, L., Hodges, P., 2003. Missing mollusks: Field testing taphonomic loss in
985 the Mesozoic through early large-scale aragonite dissolution. *Geology* **31**(3), 211–214.
- 986 Xu, H., Zhang, Y., Yuan, D., Shen, S., 2022. Quantitative palaeobiogeography of the
987 Kungurian–Roadian brachiopod faunas in the Tethys: Implications of allometric drifting of
988 Cimmerian blocks and opening of the Meso-Tethys Ocean. *Palaeogeogr. Palaeoclimatol.
989 Palaeoecol.* **601**, 111078.
- 990 Yang, S., Hu, W., Fan, J., Deng, Y., 2022. New geochemical identification fingerprints of

If you encounter any issues, please feel free to reach out to Wang Ma(马旺) (wma19952022@163.com).

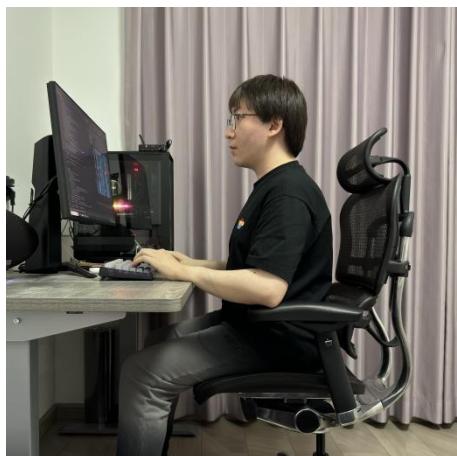
- 991 volcanism during the Ordovician-Silurian transition and its implications for biological and
992 environmental evolution, *Earth-Sci. Rev.* **228**, 104016.
993 Zhou, X., Sun, L., 2023. Factors controlling the formation and evolution of source rocks in the
994 Shahezi Formation, Xujiaweizi fault depression, Songliao Basin. *Energy Geoscience* **4**(2),
995 100140.

996 Developer



997
998 Shilong Guo (郭世龙)
999 PhD candidate at Capital Normal University

1000
1001 I am engaged in the quantitative description of the
1002 ecological succession in the Yanliao to Jehol
1003 community networks. During the use of this software, if
1004 you encounter any ecological issues, please contact
1005 me at 13820113071@163.com.
1006
1007
1008



1009
1010 Wang Ma (马旺)
1011 Bioinformatics engineer, Freshwind Biotechnology
1012 (Tianjin) Limited Company
1013
1014
1015
1016
1017
1018
1019
1020

I am currently engaged in bioinformatics research and developed the code of TaphonomeAnalyst 2.0 during my spare time. Please contact me if you have any problems with software operation at wma19952022@163.com.