

1 | Supplementary Materials 1

2 | The User Guide of TaphonomeAnalyst 2.0

Sampling process



3 | **Code availability:** The code of TaphonomeAnalyst 2.0 can be downloaded from

4 | <https://github.com/wma1995/TaphonomeAnalyst2>

5 | Development environments: Python 3.9.21 and R 4.1.3

6 | **Article:** TaphonomeAnalyst 2.0: integrative analysis software of taphocoenosis co-occurrence
7 | 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 VII)	22
21	Species correlation semi-matrix graphics (Module VIII)	24
22	Correlational Network Visualization (Module IX)	25
23	IX. 1 Overview	25
24	IX.2 Correlation	25
25	IX. 3 Louvain algorithm	27
26	Comparison of networks (Module X)	29
27	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		

39 Introduction

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

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

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

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

- 81 (1) Integrates operational taxonomic unit (OTU) abundance with geochemical data for joint
 82 analysis.
 83 (2) Adds capacity for deducing synergies between biological differences and multiple
 84 geochemical factors.
 85 (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.	✗	✓

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

87

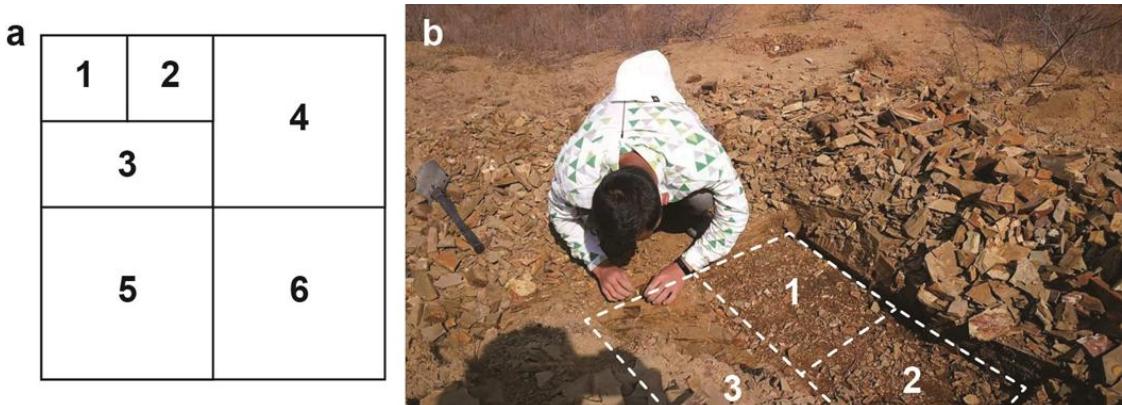
88 **The basic idea behind the software**

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

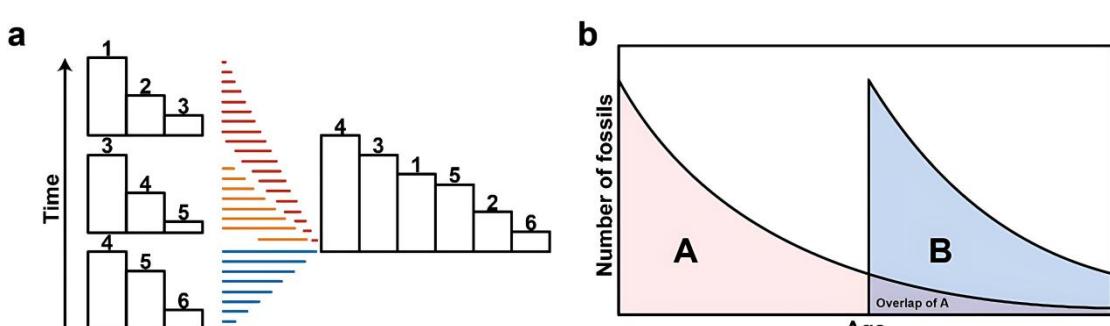
116 **Data input**

117 Fossil form a sample layer often from several centimeters of strata typically represent minimal
118 deposition spans, such as thousands of years, at least for most non-catastrophic depositional
119 environments. Fürsich and Aberhan (1990) used the term, time averaging, to describe the
120 minute chronological discrepancies existing within the same fossil layer that result from
121 bioturbation and the compression of strata. In paleoecological studies, establishing a definitive
122 temporal context for organismal burial remains a significant challenge. This limitation becomes
123 evident when considering that most fossil assemblages undergo temporal averaging
124 processes, incorporating remains from disparate communities or successive generations of
125 the same community over extended periods of time (Dhariwal et al., 2017; Karr & Clapham,
126 2015; Wing et al., 1992; Wright et al., 2003). Consequently, specimens recovered from
127 identical stratigraphic layers may represent organisms that perished and became interred at

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



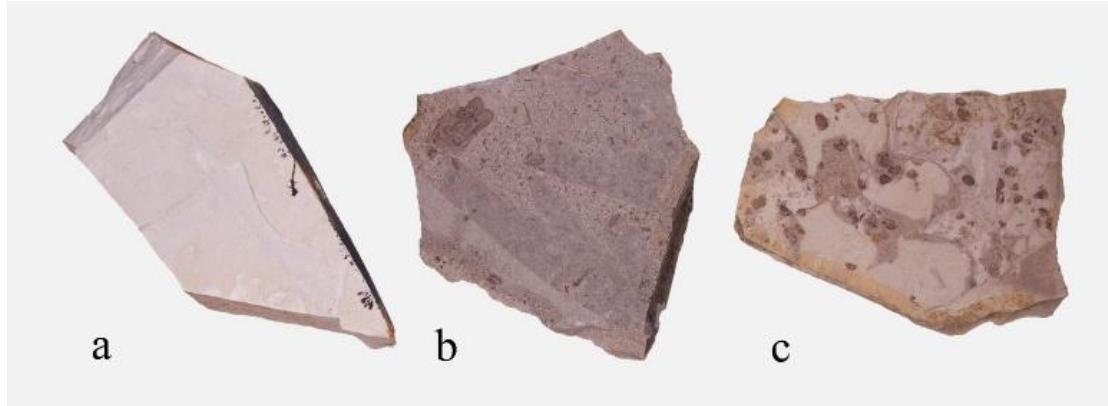
135 **Nest sampling.** **a**, Schematic diagram of the nest sampling method. The numbers represent
 136 the steps of plots excavation. **b**, a fieldwork of nest sampling.
 137



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

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

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



160
161 **Typical lithologies of the fossil samples. a,** Lithologies of Beigou. **b,** Lithologies of
162 Donggou. **c,** Lithologies of Donggouli. Our fossil example sets were collected from the latest
163 Middle Jurassic Jiulongshan Formation near Daohugou Village, in Ningcheng County of Inner
164 Mongolia, China. Based on the stratigraphic interfingering of lacustrine and a volcaniclastic
165 apron facies, the sedimentary environment of Jiulongshan Formation is interpreted as several
166 fluctuating, shallow lakes with distant volcanogenic sedimentary input. Radiometric evidence
167 shows that the depositional age of the Jiulongshan Formation is 165–164 Ma and is time
168 equivalent to the Haifanggou Formation in Liaoning Province. All fossil samples should be
169 collected from the same stratigraphic layer without significant changes in the depositional
170 environment within, to ensure stable geochemical data, such as specimens a–c above.
171

172 The recommended output format is PDF, which is more suitable for editing vector elements.
173 Common parameters:

```
174 parent_parser = argparse.ArgumentParser(add_help=False)
175 parent_parser.add_argument('--input', type=str, required=True, help='Absolute or relative path
176 file.\t[e.g. "./data.xlsx"]')
177 parent_parser.add_argument('--format', type=str, default='pdf', choices=['png', 'svg', 'pdf'],
178 help='Output format.(default: %(default)s)')
179 parser = argparse.ArgumentParser(description='A comprehensive visual software for study
180 taphonome.')
181 parser.add_argument('-v', '--version', action='version', version='TaphonomeAnalyst 2.0')
182 subparsers = parser.add_subparsers(help='commands')
183
```

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 Dp001/0001	Conchostraca	Triolventidae	Triolventa	Triolventa haifanggouensis	A	

184

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

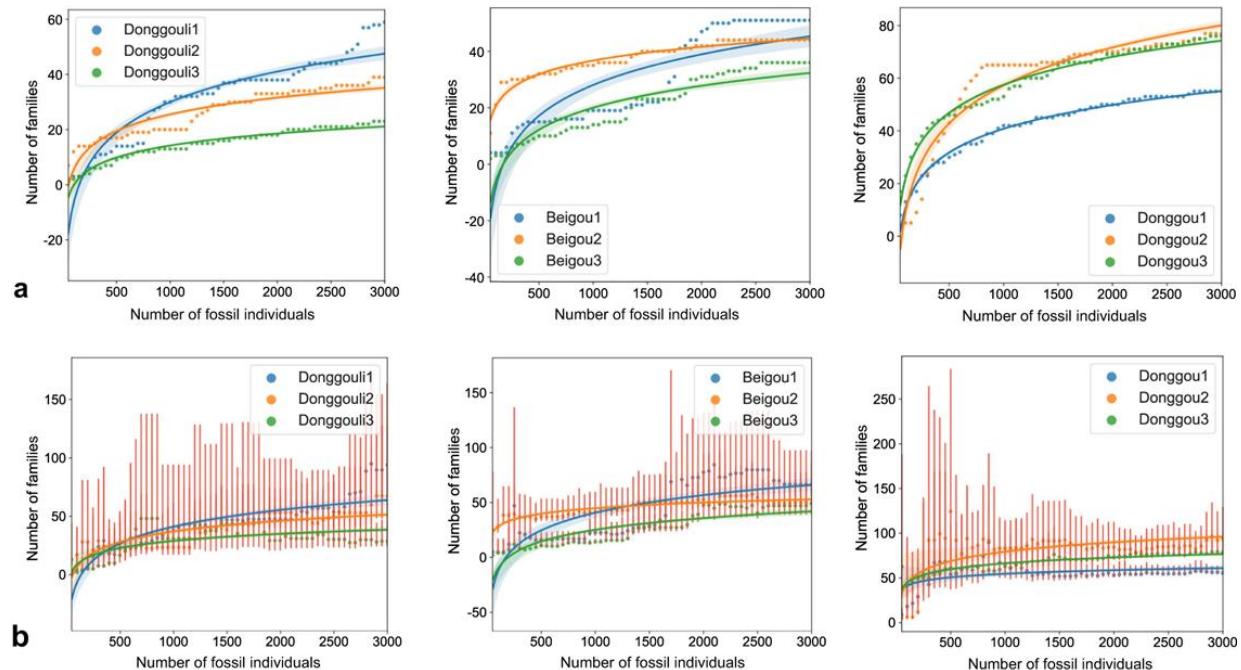
190

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

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

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

coverage, which is more reasonable. Nonetheless, due to the difference between the buried community and the present-day community, the definition of the abundance of rare species needs to be considered.



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

The Chao 1 estimator is:

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

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

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 rarity
 239 threshold-often set at 10 individuals. However, in our experience with fieldwork, a threshold of
 240 10 individuals may be excessive. Our software allows the user to set the rare species
 241 threshold when all samples are pooled. S_{rare} is the number of rare OTUs (with less than or
 242 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 $\gamma_{ace}^2 = \max \left[\frac{\frac{S_{rare}}{C_{ace}} - \frac{\sum_{i=1}^{10} i(i-1)F_i}{(N_{rare})(N_{rare}-1)}}{1, 0} \right]$ (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', 'family',  

254 '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 family  

264 --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 names.\t[e.g.  

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

276 chao_parser.add_argument('--output', type=str, default='./chao', help='Absolute path or  

277 relative path and filename.(default: %(default)s)')
```

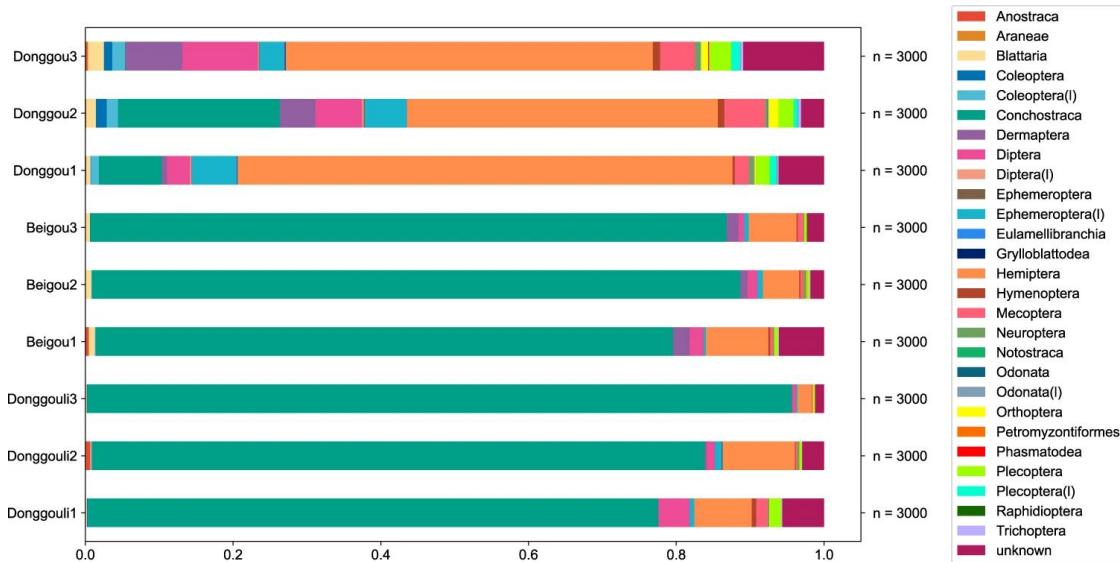
```

279 chao_parser.set_defaults(func=chao)
280 Command line:
281 python ./TaphonomistAnalyst2.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', 'genera',
293 '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
299 threshold.(default: %(default)s)')
300 ace_parser.set_defaults (func=ace)
301 Command line:
302 python ./TaphonomistAnalyst2.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-Hoffman,
313 2007; McNamara et al., 2012). To some degree, these abundances can offer insights into the
314 varying source distances and trophic levels of the species present in the fossil community.



315

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

332

333 Users have the flexibility to set the taxonomic level.

334 **Parameters:**

```
335 abundplots_parser = subparsers.add_parser(name='m2abundplots', parents=[parent_parser],  

  336 help='Abundance-sampling plots. (Module II)\t[barh]')  

  337 abundplots_parser.add_argument('--level', type=str, default='order', choices=['order', 'family',  

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

  339 abundplots_parser.add_argument('--output', type=str, default='./abundplots', help='Absolute  

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

  341 abundplots_parser.set_defaults(func=abundplots)
```

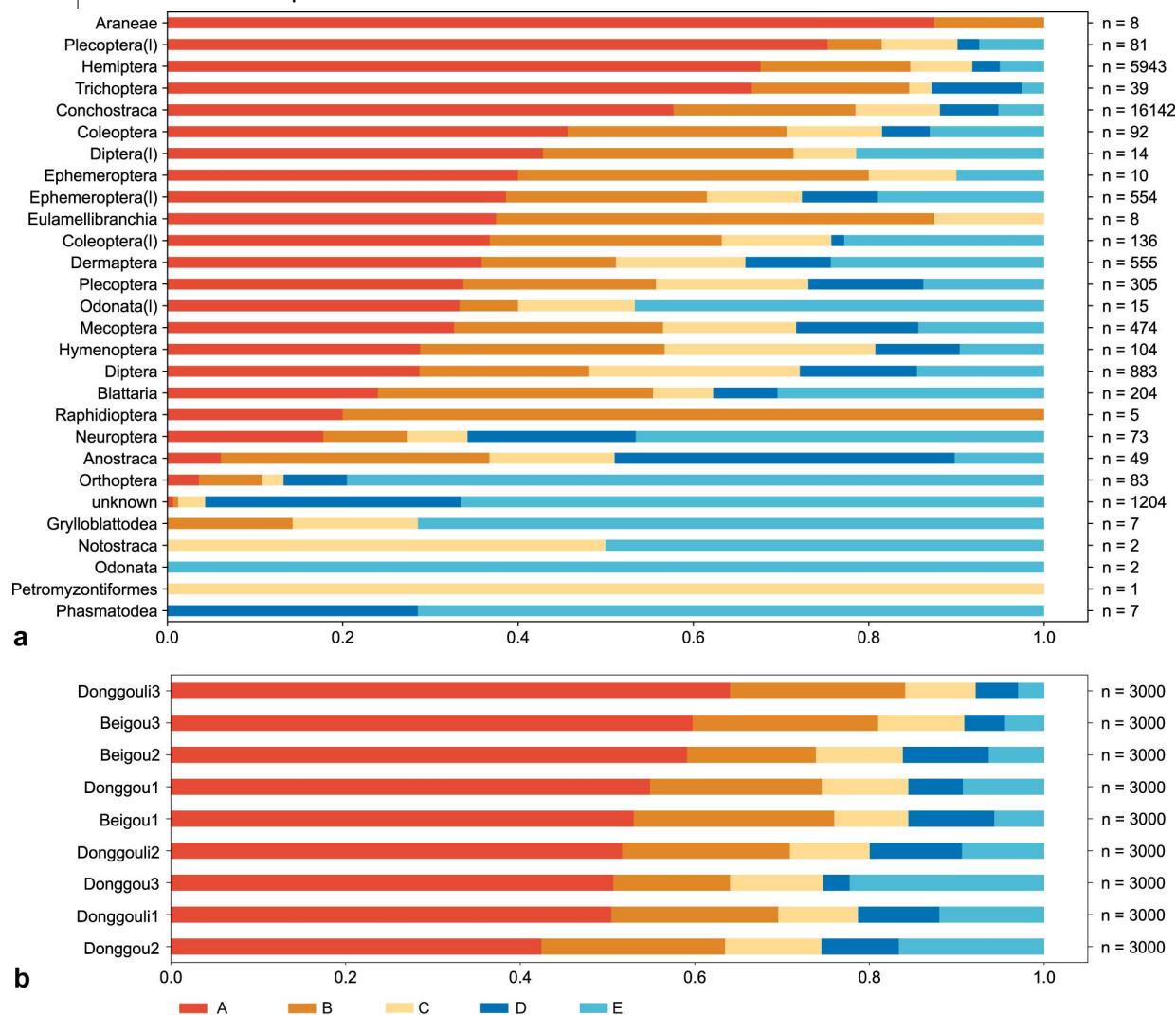
342 **Command line:**

```
343 python ./TaphonomAnalyst2.py m2abundplots --input "./Supplementary material2.xlsx" --level  

  344 order
```

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

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



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

363
364 TaphonomicAnalyst 2.0 offers the capability to selectively output the preservation levels of
365 different OTUs (operational taxonomic units) or the varying degrees of preservation across
366 different sample plots. From an inventory of 27,000 individual fossils, we found that aquatic
367 organisms are well-preserved, mostly at a high level of body completeness. Although it cannot
368 be determined whether most *T. haifanggouensis* are articulated due to their burial position;
369 nevertheless, 78.4% are observed to have a complete chitinous shell and evident growth
370 bands. In addition, >87.8% of *Y. chinensis* are well-preserved, displaying entire bodies and
371 swimmeret appendages while 79.6% of mayfly nymphs have preserved gills and cerci.
372 Furthermore, beetles that have hardened wing covers, true bugs with leathery tegmina, and
373 cockroaches and earwigs that have highly keratinized bodies also are preferentially preserved.
374 By contrast, insects with softer body surfaces and more delicate wings, such as nematoceran
375 flies, stoneflies, mayflies, and katydids, are less well-preserved. Generally, the distance
376 between the insect-occupying habitat and a water body of eventual burial is one of several
377 decisive factors in the preservation of such insects. Trichoptera adults (84.6%),
378 Ephemeroptera adults (80.0%), and Plecoptera adults (48.1%), whose larvae and naiads live
379 in water bodies, have preservation grades of A or B. Mecoptera, Hymenoptera (sawflies and
380 wasps), Diptera, and Neuroptera (lacewings), which live in humid environments, have 56.5%,
381 56.7%, 48.1%, and 27.3% scores, respectively, at A or B grades. By contrast, Orthoptera
382 (grasshoppers and katydids), Phasmatodea (stick insects), and Grylloblattodea (ice crawlers)
383 have the least degree of preservation, consisting of 0–10% scores at A or B grades. By
384 comparing the taphonomic grades from the nine plots, we found that the differences were
385 small. About 70% of the individuals are preserved at the A and B grades. These observations
386 show that most fossils are buried in an autochthonous or sub-autochthonous manner, which is
387 an ideal context for sampling.

388
389 Proportion of taphonomic grades (A–E) of the taxa. Users need to set the taxonomic level
390 themselves.

391 **Parameters:**

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

399

400 **Command line:**

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

403 **Parameters:**

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

```
407 groups.(Recommend to group the plots by different aquatic and terrestrial environments)\t[e.g.  
408 "plotA1/plotB2,plotB1/plotA2,plotC1/plotC2"]'  
409 TGplots_parser.add_argument('--output', type=str, default='./TGplots', help='Absolute path or  
410 relative path and filename.(default: %(default)s)')  
411 TGplots_parser.set_defaults(func=TGplots)  
412  
413 Command line:  
414 Default sort  
415 python ./TaphonomistAnalyst2.py m3plots --input "./Supplementary material2.xlsx"  
416  
417 Custom sort  
418 python ./TaphonomistAnalyst2.py m3plots --input "./Supplementary material2.xlsx" --groups  
419 Donggou1/Donggou2/Donggou3,Donggouli1/Donggouli2/Donggouli3/Beigou1/Beigou2/Beigo  
420 u3  
421
```

The standard of taphonomic grades

Taphonomic grades	Preservation	Clam shrimps	Other arthropods	Vertebrates
A		Shell edge >90% preserved. Growth bands are fully clear.	>90% preserved. Body articulated, wing veins visible and almost complete.	Body and limbs are complete and articulated.
B		Shell edge >70% preserved. Growth bands are almost clear.	80–90% preserved. Body almost complete, including head, thorax, abdomen and thoracic appendages, details such as antennae or cerci lost.	70–80% torso and limbs are complete. Partial joint displacement.
C		Shell edge >60% preserved. Growth bands are partially clear.	60–80% preserved. Body deformed, at least one of six legs lost.	60–70% torso preserved.
D		Shell edge >50% preserved.	30–60% preserved. Wings disarticulated, remains of head, thorax and abdomen preserved.	Torso with missing tail or head.
E		Shell fragments.	<30% preserved. High disarticulated body, isolated structures such as single legs, abdomen and/or wings preserved.	Scattered bones.

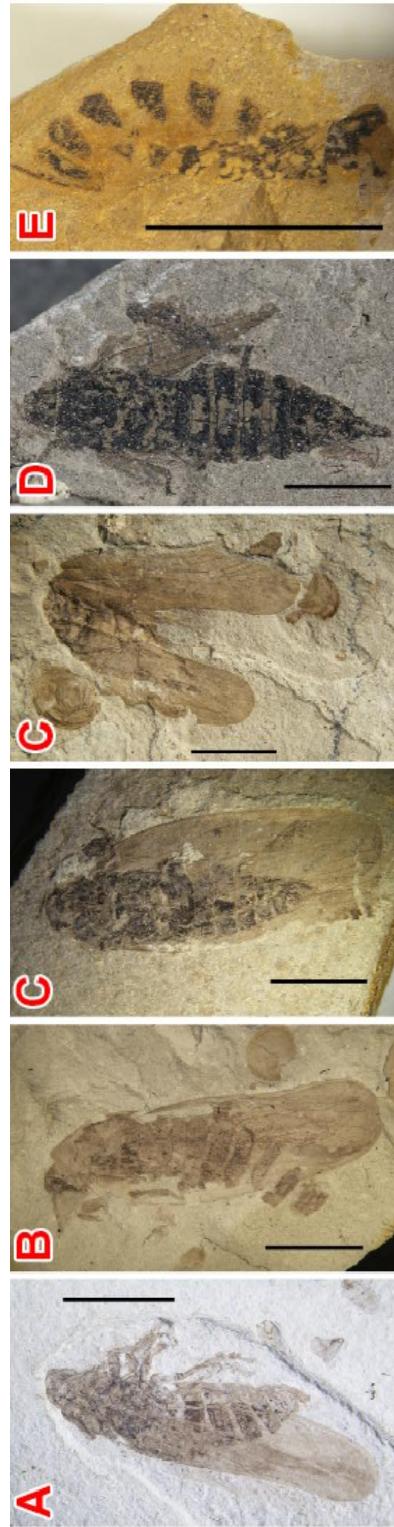


Table 2. Definition of taphonomic grades. The taphonomic grade categorizes the preservational quality of fossils into five levels, ranging from A to E. Level A signifies the highest quality of preservation, whereas level E denotes the poorest quality of preservation.

424

Taphonomic environment analysis (Module IV)

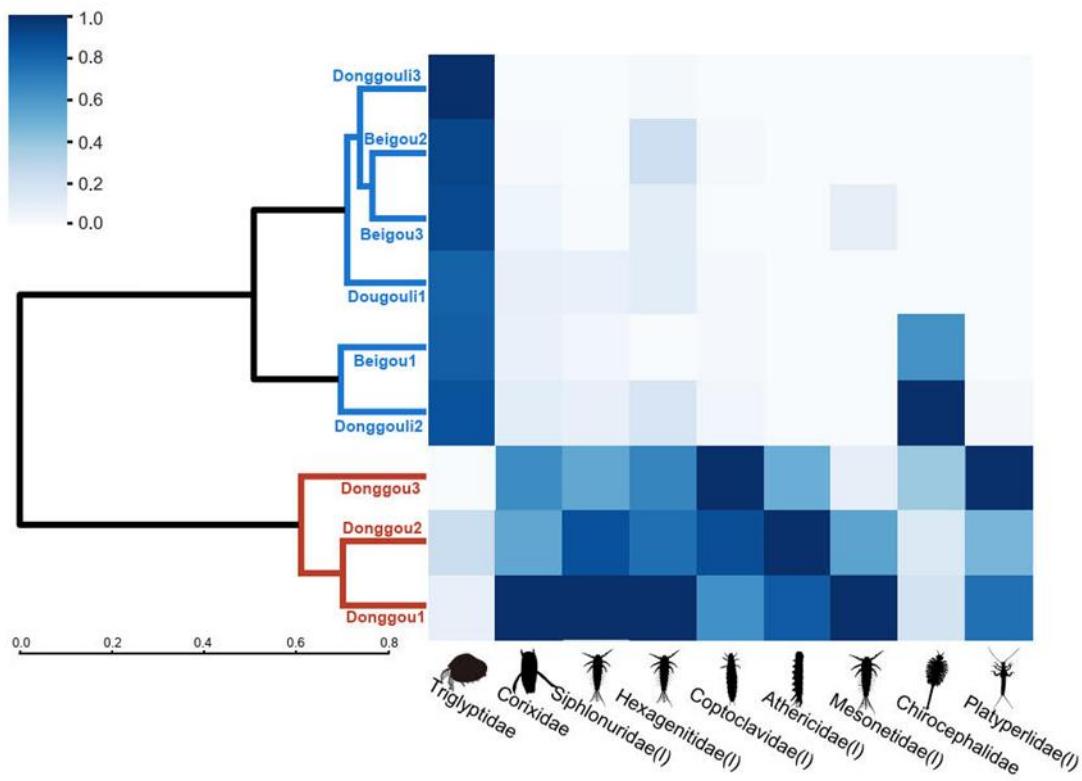
425 The primary aim of this module is to contrast and display the variations in species abundance
 426 across different sampling plots, and subsequent categorization of these samples into distinct
 427 groups. Since aquatic organisms typically undergo minimal transport during fossilization, their
 428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

438 The average assigns clustering method is:

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

440 and the Bray-Curtis distance is:

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



442

443

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

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

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

456 Parameters:

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

468

469 **Command line:**

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

477

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

482 **Parameters:**

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

```

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

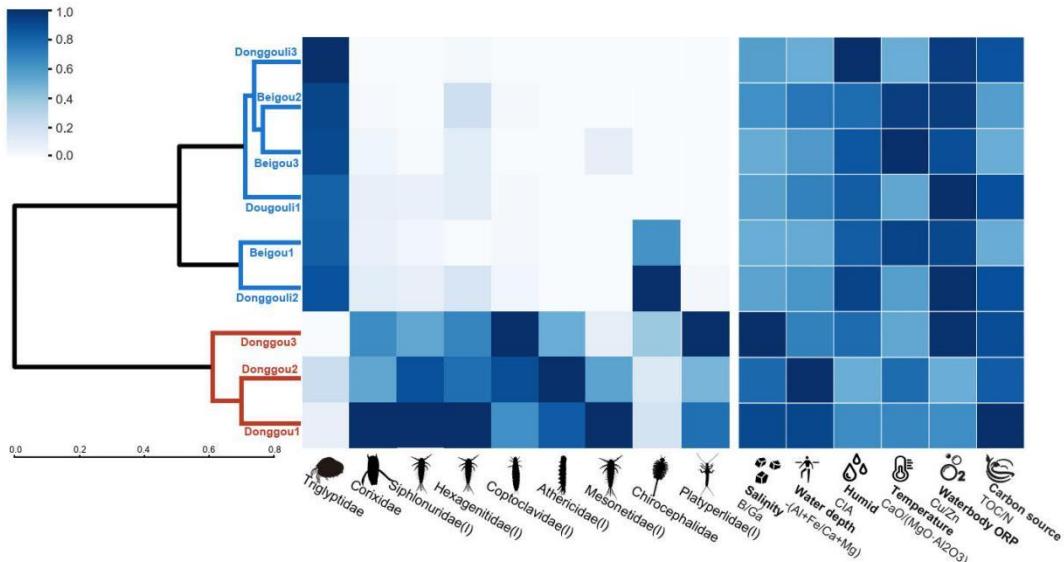
```

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

499 The environmental variables that shaped ancient ecosystems are not directly measurable and
500 often necessitate the extrapolation of geochemical data for accurate interpretation. The ratios
501 of elements and oxides have been widely used in ancient environmental studies to gain
502 insights into salinity, temperature, water depth, humidity, intensity of volcanic activities, and
503 other factors (Bai et al., 2020; Chen, Wan, 1999; Chen et al., 1999; Feng et al., 2003; Hou et al.,
504 2018; Hou et al., 2023; Fu et al., 2018; Harnois, 1988; Nesbitt, Young, 1982; McLennan, 1993;
505 Stanistreet et al., 2020; Swain et al., 2022; Wang et al., 2023; Yang et al., 2022; Zhou and Sun,
506 2023). Beyond serving as final resting grounds for fossilized remains, ancient water bodies act
507 as primary archives for geochemical information, frequently encapsulating environmental
508 contexts crucial to the majority of terrestrial organisms inhabiting aquatic habitats to littoral
509 zones. This module provides a geochemical heat map of different sampling plots. This figure
510 can be automatically combined with the sedimentary environment cluster tree and the heat
511 map distribution of aquatic organisms. It can simultaneously reflect the distribution of biological
512 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

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



514

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

519 Parameters:

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

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

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

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

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

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

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

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

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

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

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

531

532 **Command line:**

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

534 "Daohugouectes primitinus(I),Tritypta haifanggouensis,Tritypta  

535 haifanggouensis,Yanliaocorixa chinensis,Karataviella popovi,Samarura gigantea(I),Anisoptera  

536 fam. gen. sp1.(I),Platyperla platypoda(I),Ferganoconcha sibirica,Qiyia jurassica(I),Mesomyzon  

537 sp1.,Triops sp1.,Chirocephalidae gen. sp1.,Eurythoracalis mirabilis(I),Shantous  

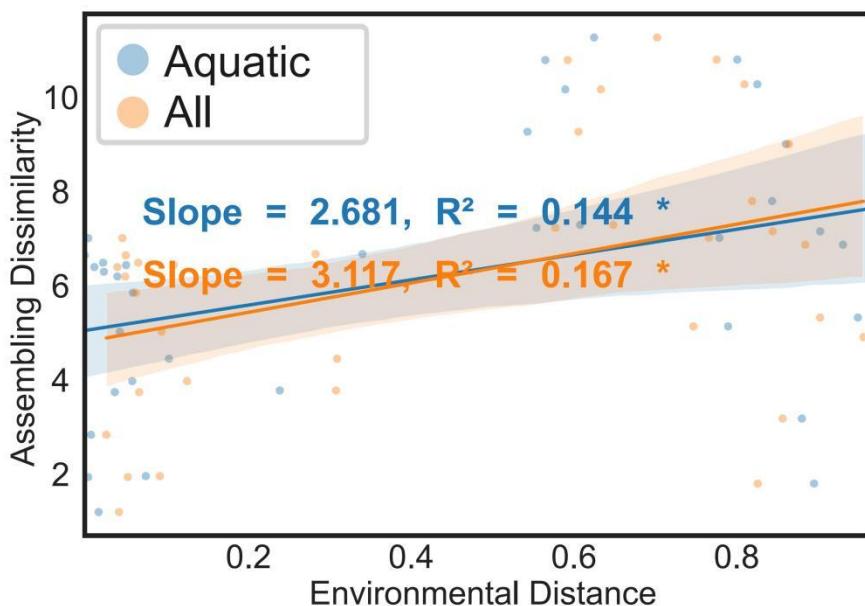
538 lacustris(I),Foliomimus latus(I),Furvoneta viriosus(I),Furvoneta raukus(I),Mesobaetis  

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

540

Assembling dissimilarity – environmental distance 541 test (Module VI)

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



549

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

556

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

559 **Parameters:**

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

```

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

```

583 Mantel Test between species abundance and 584 ecological environmental variables (Module VII)

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

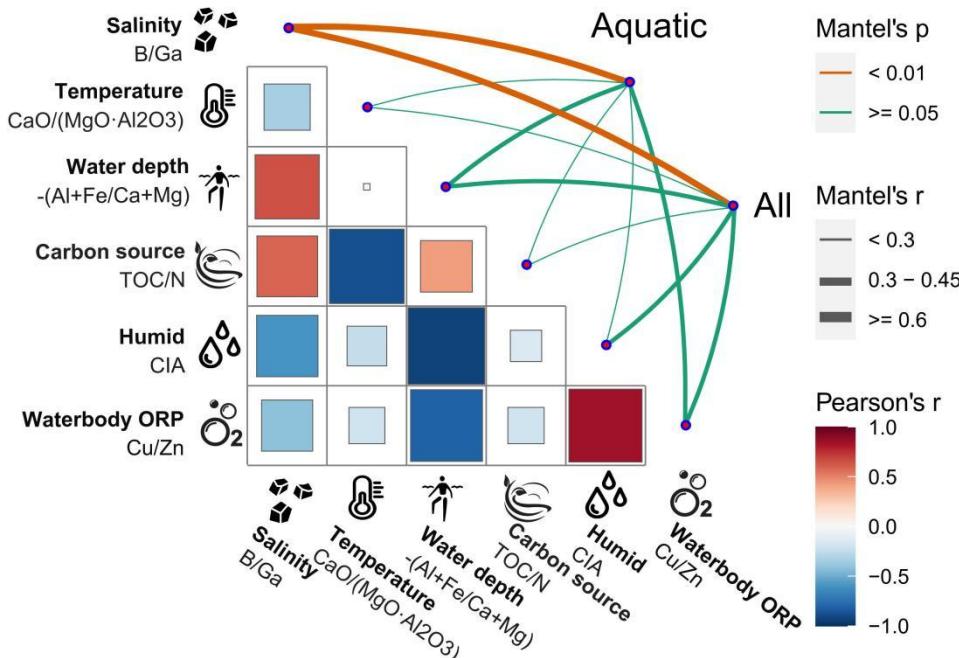
590 Since its inception, the Mantel Test has been integral to understanding the rapid evolution and
591 application across diverse scientific domains, notably in microbial community ecology. In the
592 realm of paleoecology, the Mantel Test serves as an invaluable tool for probing the
593 connections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

598 The Bray-Curtis distance is:

$$599 d_{(u,v)} = \frac{\sum_i |u_i - v_i|}{\sum_i (u_i + v_i)} \quad (4)$$

600 and the Euclidean distance is:

$$602 d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (5)$$



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

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

614 **Parameters:**

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

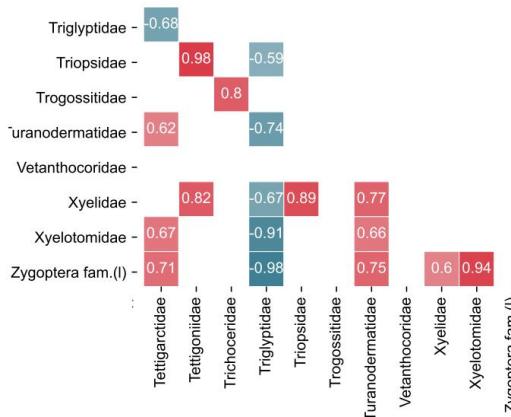
```
630 mantel_parser.add_argument('--output', type=str, default='./mantel', help='Absolute path or  
631 relative path and filename.(default: %(default)s)'  
632 mantel_parser.set_defaults(func=mantel)  
633
```

634 **Command line:**

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

643 **Species correlation semi-matrix graphics (Module VIII)**

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



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

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

```
652
```

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

```
655 Parameters:
```

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

```

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

```

671 Correlational Network Visualization (Module IX)

672 IX, 1 Overview

673 A network characterized by nodes and links is a composition of both elements. In such a
674 network, the nodes symbolize taxa found within sampling plots, while the edges represent the
675 taphonomic co-occurrences among these taxa. Animals inhabiting the same microenvironment
676 also demonstrate co-occurrence during the process of fossilization. Therefore, the
677 phenomenon of taphonomic co-occurrence partly can be explained by the extent of habitat
678 overlap between two taxa. In constructing the network, two crucial considerations must be
679 addressed: the methodology for correlation calculation and division of 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 compositional
683 data) coefficients for network visualization. It also enables users to define their own filters,
684 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 of
690 correlation between microbial abundances by log-transforming the observed abundance data
691 and using the variance of log ratios to correct for biases. The method also employs a bootstrap
692 procedure to calculate P-values, allowing for assessment of the statistical significance of the
693 correlations. However, its performance is constrained when computing interaction networks for
694 high dimensional datasets. Typically, microbial network studies require a minimum of 3000
695 samples within the total dataset, with a OTUs diversity not exceeding 800.

696 SparCC is based on the log-ratio transformation:

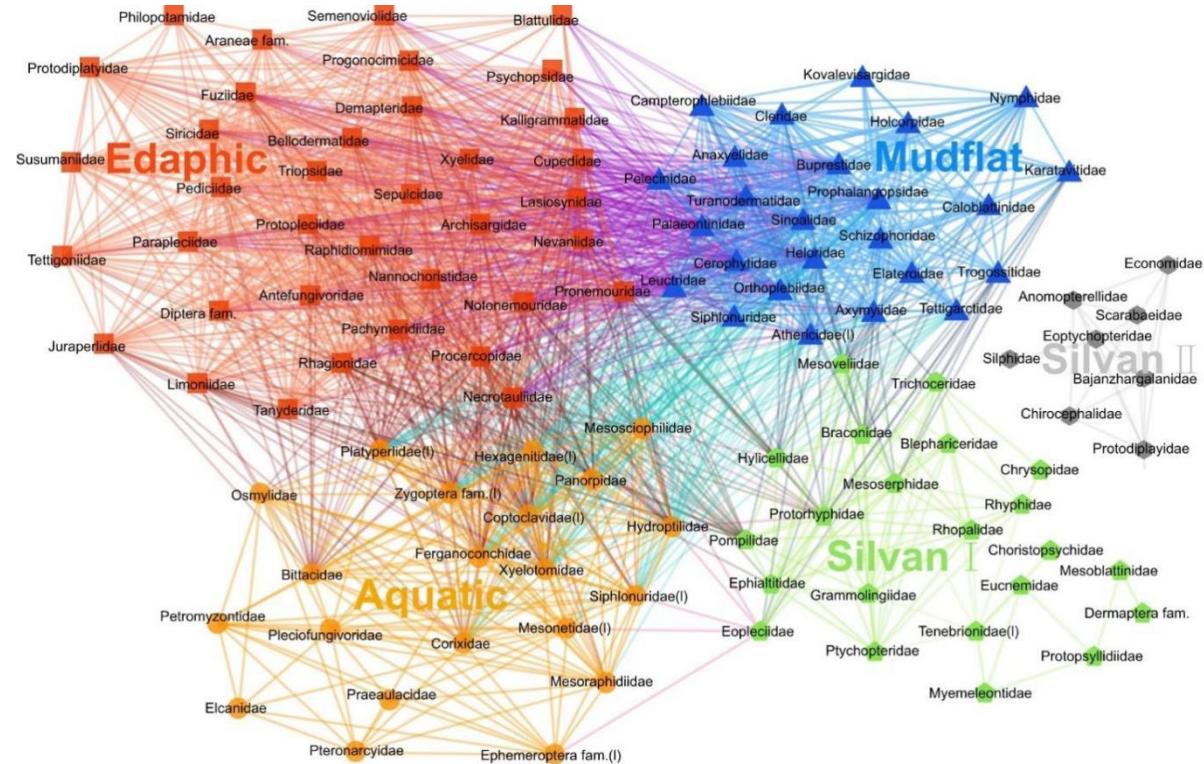
$$697 y_{ij} = \log \frac{x_i}{x_j} = \log x_i - \log x_j \quad (6)$$

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

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

701 $t_{ij} = \text{Var}\left[\frac{x_j}{x_i}\right] = \text{Var}[y_{ij}] \quad (7)$

702 When OUT i and j are absolutely correlated and their ratio is constant; consequently, $t_{ij}=0$
 703 and $t_{ij} = w_i^2 + w_j^2 - 2\rho_{ij}w_iw_j$
 704 w_i^2, w_j^2 : The variances of the log-transformed basis abundances OUT i and j. ρ_{ij} : the
 705 correlation between them OUT i and j.



706 **Network visualization.** Points of the same color represent the same module, indicating that
 707 these organisms likely inhabit the same environment. The output image is in PDF format,
 708 allowing users to adjust the font size in a PDF editor to meet publishing requirements. Based
 709 on randomization tests, this threefold community structure is statistically significant (Pearson
 710 correlation > 0 , $P < 0.1$). This algorithm segregates families into five modules, which also
 711 means five environments. The first is an aquatic assemblage (orange) that includes water
 712 boatman (Corixidae); mayfly naiads of Mesonetidae, Siphlonuridae, and Hexagenitidae,
 713 naiads of stoneflies and dragonflies, *Daohugouneutes primitinus*, *Ferganoconcha sibirica* and
 714 a lamprey genus in the Petromyzontidae. This assemblage represents a deep-water
 715 environment. Second, is a mudflat assemblage (blue), that includes adults of soldier flies
 716 (Kovalevisargidae), axymyiids (Axymyiidae), snipe flies (Rhagionidae), primitive crane flies
 717 (Tanyderidae), sinoalids (Sinoalidae); and adults of aquatic insects such as primitive minnow
 718 mayflies (Siphlonuridae) and rolled-winged stoneflies (Leuctridae). This assemblage is typical
 719 of wet environments colonized by herbaceous plants adjacent to water bodies. Third, is the
 720 edaphic assemblage (red) that consists generally of cockroaches of the Fuziidae and
 721 Blattulidae; earwigs of the Bellodermatidae, Dermapteridae, and Turanodermatidae; and a
 722 beetle assigned to Lasiosynidae. This assemblage expressed an affinity for an above-ground

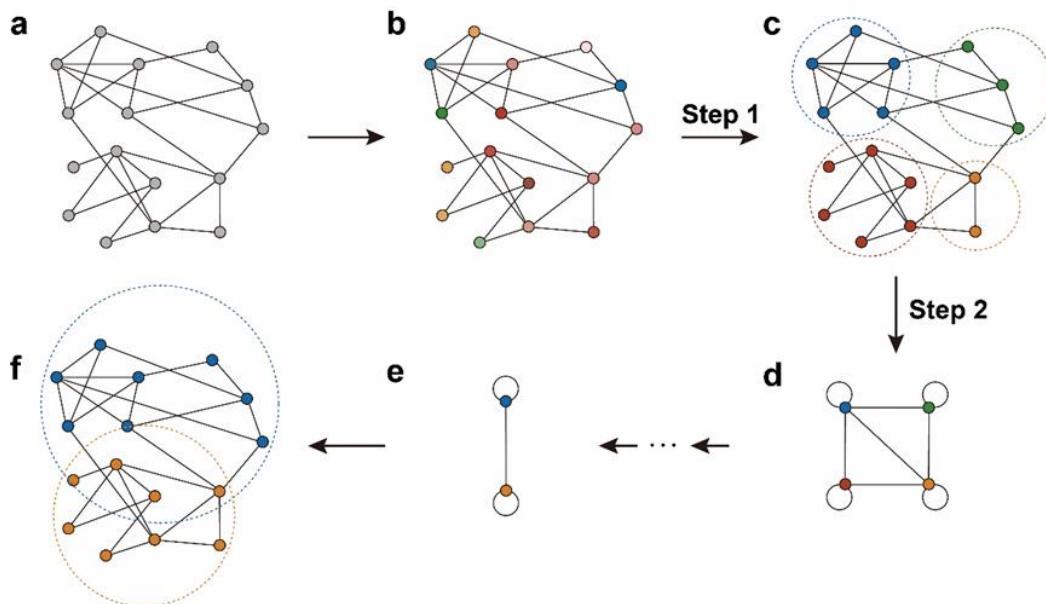
723 environment of humus and litter of ambient vegetation. Fourth, is the silvan assemblage (bright
 724 green) that includes a green lacewing (Chrysopidae), a wood gnat (Rhyphidae), an antlion
 725 (Myrmeleontidae), a lacewing (Grammolongiidae), and a primitive, cicada-like form
 726 (Protopsylliidae). This assemblage shows a preference for a relatively dry environment. A
 727 small assemblage (grey), including scarab beetles (Scarabaeidae), caddisflies (Economidae),
 728 and primitive crane flies (Eoptychopteridae), probably inhabited forests.

729 IX. 3 Louvain algorithm

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

$$738 Q = \frac{1}{2m} \sum_{v\omega} \left[A_{v\omega} - \frac{k_v k_\omega}{2m} \right] \delta(c_v, c_\omega) \quad (8)$$

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

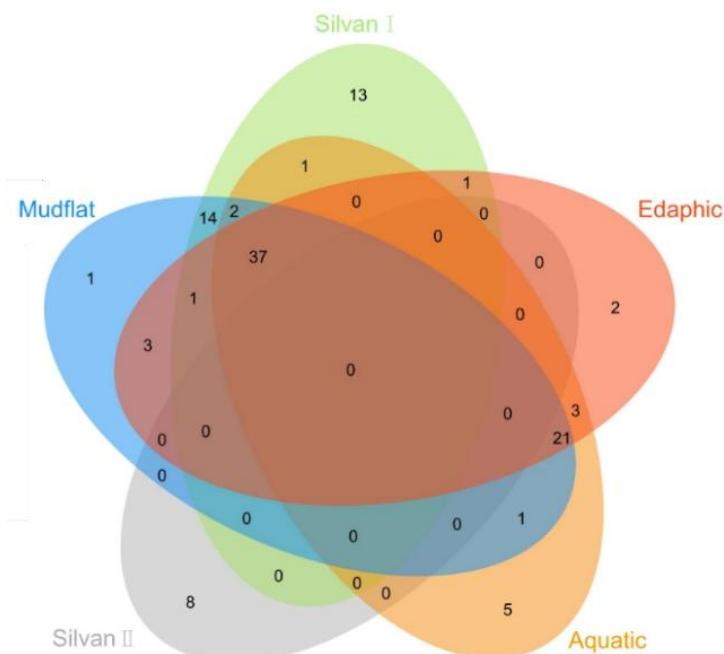


742 **The operation procedure of Louvain algorithm.** **a**, The network is constructed based on
 743 correlation matrix. **b, c**, Step one: each node in the network is assigned to its individual module.
 744 The Louvain algorithm examines the increase in modularity if neighboring nodes of node u are
 745 assigned to the same module. If there is no increase in modularity results when neighboring
 746 nodes joined (i.e., the gain in modularity is zero or negative), then node-neighboring nodes
 747 remain in their original module. This procedure is repeatedly carried out until no further
 748 enhancement in modularity can be achieved by altering the community assignments of any
 749 nodes, at which point the first phase is concluded. **c, d**, Step two: after the algorithm divides
 27

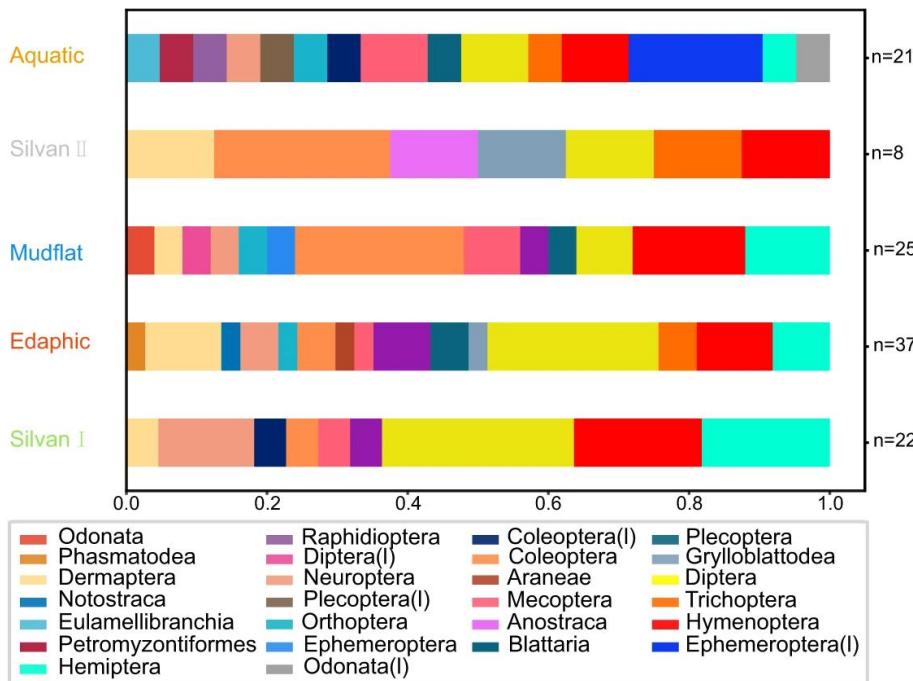
750 the modules in the first round, each module is merged into a large self-looping node. The
 751 weight of the self-looping node is the sum of the link weights of all nodes in the original module.
 752 The newly merged nodes's weight is the sum of the link weights of all nodes in the original
 753 module. **d, e**, The subsequent calculation method is the same as step one. **f**, The final output
 754 of the resulting graph. The increase in community modularity can be computed using the
 755 following formula:

$$756 \Delta Q = \left[\frac{\sum_{in} + k_{v,in}}{2m} - \left(\frac{\sum_{tot} + k_v}{2m} \right)^2 \right] - \left[\frac{\sum_{in}}{2m} - \left(\frac{\sum_{tot}}{2m} \right)^2 - \left(\frac{k_v}{2m} \right)^2 \right] \quad (9)$$

757 \sum_{in} is the sum of the weights of all links in same module. \sum_{tot} is the sum of the weights of all
 758 links that are external to the module. k_i is the sum of weights of node v . $k_{v,in}$ is the sum of
 759 weights between node v to the module whereby node v is moved.



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



762

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

764 Users need to input the taxonomic level, types and strength conditions of the correlations, and
765 p-value conditions.

766 Parameters:

```
767 cooccurnet_parser = subparsers.add_parser(name='m9cooccurnet', parents=[parent_parser],  
768 help='Co-occurrence networks. (Module IX)\t[network/venn]')  
769 cooccurnet_parser.add_argument('--level', type=str, default='family', choices=['order', 'family',  
770 'genera', 'species'], help='Taxonomic level.(default: %(default)s)')  
771 cooccurnet_parser.add_argument('--corr', type=str, default='pearson', choices=['pearson',  
772 'spearman', 'kendall', 'sparcc'], help='Correlation algorithm.(default: %(default)s)')  
773 cooccurnet_parser.add_argument('--corr_coef', type=float, default=0.7, help='Minimum  
774 threshold of correlation coefficient.(default: %(default)s)')  
775 cooccurnet_parser.add_argument('--p_value', type=float, default=0.1, help='Maximum  
776 threshold of p-value.(default: %(default)s)')  
777 cooccurnet_parser.add_argument('--output', type=str, default='./cooccurnet', help='Absolute  
778 path or relative path and filename.(default: %(default)s)')  
779 cooccurnet_parser.set_defaults(func=cooccurnet)
```

780

781 Command line:

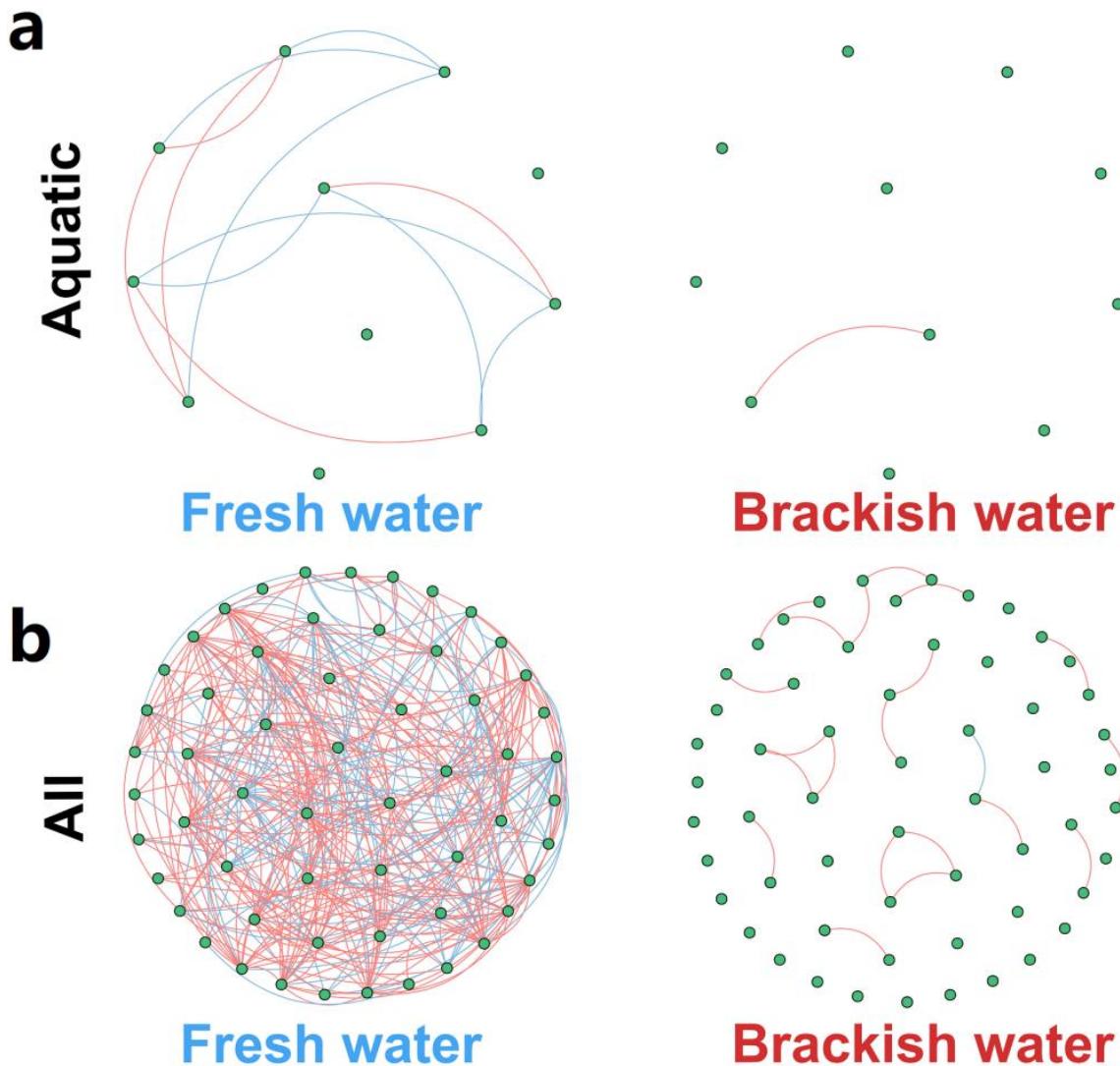
```
782 python ./TaphonomAnalyst2.py m9cooccurnet --input "./Supplementary material2.xlsx" --level  
783 family --corr pearson --corr_coef 0.7 --p_value 0.1
```

784

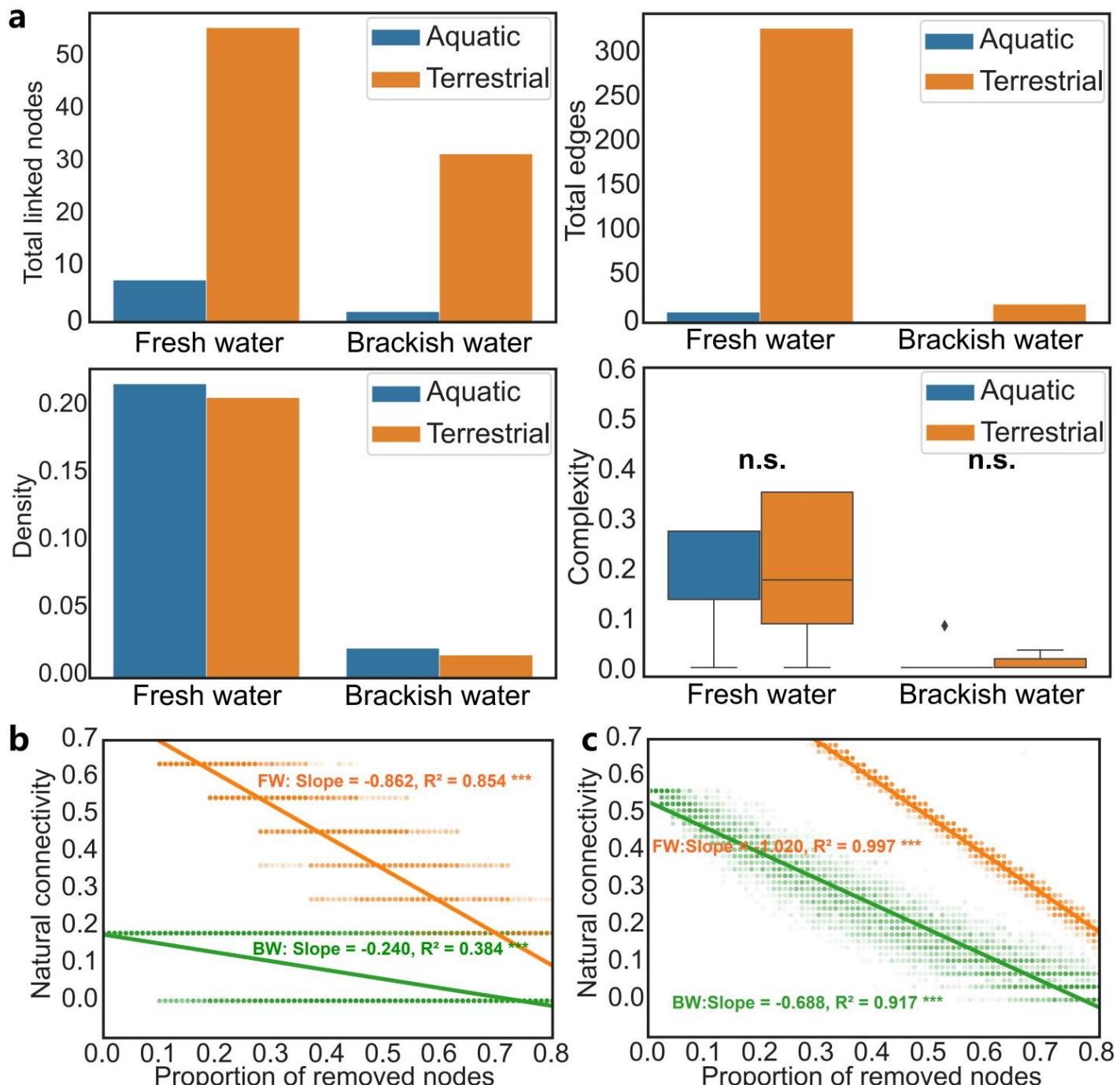
Comparison of networks (Module X)

785 This module offers the capability to compare networks under different groups of sampling plots.
786 The network set for comparison can be easy to read by using a consistent, random layout.

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



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



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

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

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

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

811 **Parameters:**

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

833 **Command line:**

```
834 python ./TaphonomeAnalyst2.py m10netVC --input "./Supplementary material2.xlsx" --aquatic  
835 "Daohugouectes primitinus(I),Triglypta haifanggouensis,Triglypta  
836 haifanggouensis,Yanliaocorixa chinensis,Karataviella popovi,Samarura gigantea(I),Anisoptera  
837 fam. gen. sp1.(I),Platyperla platypoda(I),Ferganoconcha sibirica,Qiyia jurassica(I),Mesomyzon  
838 sp1.,Triops sp1.,Chirocephalidae gen. sp1.,Eurythoracalis mirabilis(I),Shantous  
839 lacustris(I),Foliomimus latus(I),Furvoneta viriosus(I),Furvoneta raukus(I),Mesobaetis  
840 sibirica(I),Clavineta eximia(I)" --level_aquatic species --level_terrestrial family --groups  
841 Donggou1/Donggou2/Donggou3,Donggouli1/Donggouli2/Donggouli3/Beigou1/Beigou2/Beigo  
842 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. $d(G) = \frac{2L}{N(N-1)} \quad (10)$	The complexity of a symbiotic relationship.
	N: Total nodes. L: Total edges.	
Complexity	The complexity of the network. $C_i = \frac{d_i}{n} \quad (11)$ 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. $\deg(v) = \sum_{u \in V, u \neq v} A_{uv} \quad (12)$ 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

844

845

Table 3. Evaluation indicators of the network.

846

847 Frequently Asked Questions

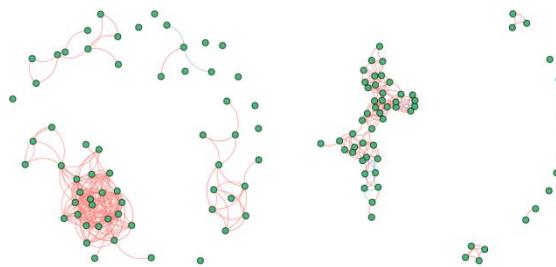
848 **The error of identical OTUs being identified as two 849 separate ones.**

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

851 **The proportion of taphonomic grades error**

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

855 **The network comparison error**



856
857 The network comparison visualization may face challenges due to an overabundance of links.
858 This plethora of connections complicates the software's ability to identify matching layouts. We
859 suggest implementing the spearman correlation coefficient or imposing stricter correlation
860 criteria to diminish the link count.

861 **The geochemical heat map error**

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

867 Future developments

868 TaphonomAnalyst 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 TaphonomAnalyst 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 TaphonomAnalyst 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 dplyr,
882 linkET, and ggplot2.

883 Acknowledgments

884 The original version of TaphonomAnalyst 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 TaphonomAnalyst
889 manuscript: Roxanne Bunker, 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.
908 <https://doi.org/10.1088/1742-5468/2008/10/P10008>.
- 909 Chao, A., 1984. Non-parametric estimation of the number of classes in a population. *Scand. J.*
910 *Stat.* **11**, 265–270.
- 911 Chao, A., Lee, S.M., 1992. Estimating the number of classes via sample coverage. *J. Am. Stat.*
912 *Assoc.* **87** (417), 210–217.
- 913 Chao, A., Shen, T. J., 2004. Nonparametric prediction in species sampling. *J. Agr. Biol. Envi. St.*
914 **9**, 253–269.
- 915 Chao, A., Yang, M.C. K., 1993. Stopping rules and estimation for recapture debugging with
916 unequal failure rates. *Biometrika* **80** (1), 193–201.
- 917 Chen, J., Wan, G., 1999. Sediment particle size distribution and its environmental significance
918 in Lake Erhai, Yunnan province. *Chin. J. Geochem.* **18**(4), 314–320.
- 919 Chen, J., Wan, G., Chen, Z., Huang, R., 1999. Chemical elements in sediments of Lake Erhai
920 and palaeoclimate evolution. *Geochimica* **28**(6), 562–570.
- 921 Dhariwal, A., Chong, J., Habib, S., King, I.L., Agellon, L.B., Xia, J., 2017. Microbiomeanalyst: a
922 web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data.
923 *Nucleic. Acids. Res.* **45** (W1), W180–W188.
- 924 Feng, L., Chu, X., Zhang, Q., 2003. CIA (chemical index of alteration) and its applications in
925 the Neoproterozoic clastic rocks. *Earth Sci. Front.* **10**(4), 539–544.
- 926 Fu, J., Li, S., Xu, L., Niu, X., 2018. Paleo-sedimentary environmental restoration and its
927 significance of Chang 7 Member of Triassic Yanchang Formation in Ordos Basin, NW
928 China. *Pet. Explor. Dev.* **45**(6), 998–1008.
- 929 Fürsich, F. T., Aberhan, M., 1990. Significance of time-averaging for palaeocommunity analysis.
930 *Lethaia* **23**, 143–152.
- 931 Guo, S., Ma, W., Tang, Y., Chen, L., Wang, Y., Cui, Y., Liang, J., Li, L., Zhuang, J., Gu, J., Li, M.,
932 Fang, H., Lin, X., Shih, C.K., Labandeira, C.C., Ren, D., 2023. A new method for
933 examining the co-occurrence network of fossil assemblages. *Commun. Biol.* **6**(1), 1102.
934 <https://doi.org/10.1038/s42003-023-05417-6>
- 935 Harnois, L., 1988. The C.I.W Index: a new chemical index of weathering. *Sediment. Geol.*
936 **55**(3), 319–322.
- 937 Hou, Q., Jin, Q., Niu, C., Zhang, R., Chen, F., Xu, J., Zhang, F. J., 2018. Distribution
938 characteristics and main controlling factors of main hydrocarbon sourcerocks in Liaodong
939 bay area. *Earth Sci.* **43**(6), 2160–2171.
- 940 Karr, J., Clapham, M., 2015. Taphonomic biases in the insect fossil record: Shifts in articulation
941 over geologic time. *Paleobiology* **41**(1), 16–32.
- 942 Kurtz, Z.D., Müller, C.L., Miraldi, E.R., Littman, D.R., Blaser, M.J., Bonneau, R.A., 2015 Sparse
943 and Computationally Robust Inference of Microbial Ecological Networks. *PLoS Comput.*
944 *Biol.* **11**(5): e1004226. <https://doi.org/10.1371/journal.pcbi.1004226>
- 945 Mantel, N. (1967). The detection of disease clustering and a generalized regression approach.
946 *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 time
962 averaging and differential preservation on macroinvertebrate species richness in death
963 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. Potts,
983 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 the
985 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.

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

- 990 Yang, S., Hu, W., Fan, J., Deng, Y., 2022. New geochemical identification fingerprints of
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.

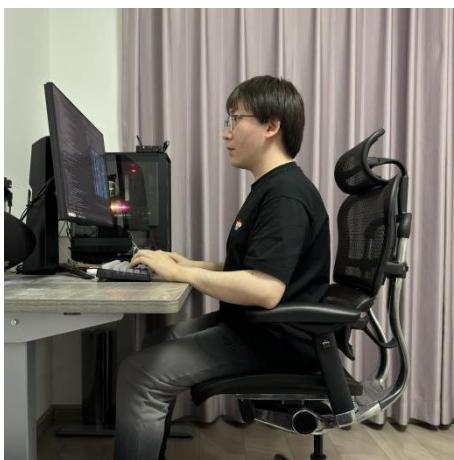
996 Developer



Shilong Guo (郭世龙)

PhD candidate at Capital Normal University

I am engaged in the quantitative description of the ecological succession in the Yanliao to Jehol community networks. During the use of this software, if you encounter any ecological issues, please contact me at 13820113071@163.com.



Wang Ma (马旺)

Bioinformatics engineer, Freshwind Biotechnology (Tianjin) Limited Company

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.