

1 Scientific drilling projects in ancient lakes: 2 Integrating geological and biological histories

3 Thomas Wilke^{a,*}, Bernd Wagner^{b,c}, Bert Van Bocxlaer^{a,d,e}, Christian Albrecht^a, Daniel Ariztegui^f,
4 Diana Delicado^a, Alexander Francke^b, Mathias Harzhauser^g, Torsten Hauffe^a, Jens Holtvoeth^h, Janna
5 Just^b, Melanie J. Leng^{i,j}, Zlatko Levkov^k, Kirsty Penkman^l, Laura Sadori^m, Alister Skinnerⁿ, Björn
6 Stelbrink^a, Hendrik Vogel^o, Frank Wesselingh^p, Thomas Wonik^q

7 ^a Department of Animal Ecology and Systematics, Justus Liebig University Giessen, Giessen, Germany

8 ^b Institute of Geology and Mineralogy, University of Cologne, Cologne, Germany

9 ^c Collaborative Research Center 806 – Our Way to Europe, University of Cologne, Cologne, Germany

10 ^d Research Unit Palaeontology and Limnology Group, Ghent University, Ghent, Belgium

11 ^e Museum für Naturkunde – Leibniz Institute for Evolution and Biodiversity, Berlin, Germany

12 ^f Department of Earth Sciences, University of Geneva, Geneva, Switzerland

13 ^g Geological-Paleontological Department, Natural History Museum Vienna, Vienna, Austria

14 ^h School of Chemistry, University of Bristol, Bristol, UK

15 ⁱ Centre for Environmental Geochemistry, School of Geography, University of Nottingham, Nottingham, UK

16 ^j NERC Isotope Geosciences Facilities, British Geological Survey, Keyworth, Nottingham, UK

17 ^k Institute of Biology, University Ss Cyril and Methodius, Skopje, Macedonia

18 ^l Department of Chemistry, University of York, York, UK

19 ^m Dipartimento di Biologia Ambientale, Università di Roma "La Sapienza", Rome, Italy

20 ⁿ ACS Core Services, Edinburgh, UK

21 ^o Institute of Geological Sciences and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland

22 ^p National Museum of Natural History, Leiden, The Netherlands

23 ^q Leibniz Institute for Applied Geophysics, Hannover, Germany

24 * Corresponding author at: Department of Animal Ecology, Justus Liebig University, Heinrich-Buff-Ring 26 (IFZ), 35392
25 Giessen, Germany. Tel.: +49 641 99 25720. E-mail address: tom.wilke@allzool.bio.uni-giessen.de (T. Wilke).

26 *Keywords:* ancient lake; long-lived lake; deep drilling; evolutionary biology; methodology;
27 paleolimnology

28	1.	Introduction
29	2.	Data and methods
30	2.1.	<i>Geological data and methods</i>
31	2.1.1.	<i>Site selection and drilling strategies</i>
32	2.1.2.	<i>Coring techniques and data recording</i>
33	2.1.3.	<i>Borehole logging</i>
34	2.1.4.	<i>Sediment-core data</i>
35	2.1.5.	<i>Age-depth models</i>
36	2.1.6.	<i>Stable isotopes</i>
37	2.2.	<i>Biological and paleontological data and methods</i>
38	2.2.1.	<i>Macrofossils</i>
39	2.2.2.	<i>Microfossils</i>
40	2.2.3.	<i>Subsurface biosphere</i>
41	2.2.4.	<i>Biomarkers</i>
42	2.2.5.	<i>Ancient amino acids</i>
43	2.2.6.	<i>Ancient DNA</i>
44	2.3.	<i>Data accessibility and storage</i>
45	3.	Integrating geological and biological data
46	3.1.	<i>Paleolimnological reconstruction</i>
47	3.3.	<i>Molecular dating</i>
48	3.4.	<i>Tracing character evolution</i>
49	3.2.	<i>Consequences of environmental change: A biological diversification perspective</i>
50	4.	Conclusions
51		Glossary
52		Acknowledgements
53		References

54 ABSTRACT

55 Sedimentary sequences in ancient or long-lived lakes can reach several thousands of meters in
56 thickness and often provide an unrivalled perspective of the lake's regional climatic, environmental,
57 and biological history. Over the last few years, deep drilling projects in ancient lakes became
58 increasingly multi- and interdisciplinary, as, among others, seismological, sedimentological,
59 biogeochemical, climatic, environmental, paleontological, and evolutionary information can be
60 obtained from sediment cores. However, these multi- and interdisciplinary projects pose several
61 challenges. The scientists involved typically approach problems from different scientific perspectives
62 and backgrounds, and setting up the program requires clear communication and the alignment of
63 interests. One of the most challenging tasks, besides the actual drilling operation, is to link diverse
64 datasets with varying resolution, data quality, and age uncertainties to answer interdisciplinary
65 questions synthetically and coherently. These problems are especially relevant when secondary data,
66 i.e., datasets obtained independently of the drilling operation, are incorporated in analyses.
67 Nonetheless, the inclusion of secondary information, such as isotopic data from fossils found in
68 outcrops or genetic data from extant species, may help to achieve synthetic answers. Recent
69 technological and methodological advances in paleolimnology are likely to increase the possibilities of
70 integrating secondary information, e.g., through molecular dating of molecular phylogenies. Some of
71 the new approaches have started to revolutionize scientific drilling in ancient lakes, but at the same
72 time, they also add a new layer of complexity to the generation and analysis of sediment core data.
73 The enhanced opportunities presented by new scientific approaches to study the paleolimnological
74 history of these lakes, therefore, come at the expense of higher logistic, communication, and analytical
75 efforts. Here we review types of data that can be obtained in ancient lake drilling projects and the
76 analytical approaches that can be applied to empirically and statistically link diverse datasets for
77 creating an integrative perspective on geological and biological data. In doing so, we highlight
78 strengths and potential weaknesses of new methods and analyses, and provide recommendations for
79 future interdisciplinary deep drilling projects.

80

81 **1. Introduction**

82 The vast majority of the world's lakes has existed or will exist for up to a few ten thousand years
83 (e.g., Brooks, 1950). Primarily due to sediment infill, they become progressively shallower and
84 subsequently vanish. Ancient or long-lived lakes, on the contrary, exist for over 100,000 years (100
85 ky), sometimes millions of years (My) (Brooks, 1950; Gorthner, 1994; Martens, 1997). They typically
86 occur in settings where sedimentation rates are low or balanced by subsidence (Cohen, 2012).
87 Accordingly, most of today's ancient lakes are oligotrophic and situated in active tectonic graben
88 settings or impact craters with low sediment supply from the catchment.

89 Because of the long-term availability of accommodation space (Jervey, 1998), sediment sequences
90 in ancient lakes can reach several hundreds to thousands of meters in thickness (e.g., Scholz et al.,
91 1993, 2011; Lindhorst et al., 2015). Lake deposits contain material that mostly derives from the lake
92 proper and the catchment area and, hence, provide an unparalleled perspective of the lake's history
93 through time (O'Sullivan, 2004). Combining the paleolimnological records from different lakes
94 permits to reconstruct continental and global environmental, and climatological histories. It is this
95 potential, captured in the often continuous lacustrine sedimentary archives, that has inspired several
96 deep drilling projects in ancient lakes (reviewed in Cohen, 2012; Fig. 1).

97 However, over the past decades, drilling operations became increasingly multidisciplinary, as data
98 bearing on physical, chemical, biochemical, and biological research questions can also be obtained
99 from sediment cores. Because of a wealth of new information, scientists from different fields, such as
100 sedimentology, climatology, geochemistry, paleolimnology, paleontology, biochemistry,
101 microbiology, evolutionary biology, physics, and modeling, currently aim to use ancient lakes as
102 paradigms to interactively look into natural phenomena from various angles, emphasizing the need for
103 truly interdisciplinary collaborations (*sensu* O'Sullivan, 2004; Birks and Birks, 2006).

104



105

106 **Fig. 1.** Map showing location of ancient lakes with a presumed age of > 1 My at which deep drilling has been
107 done (Cohen, 2012; Russell and Bijaksana, 2012).

108

109 Multidisciplinary and interdisciplinary studies enable a more holistic approach to scientific
110 problems, provide excellent opportunities for hypothesis-driven research, and are likely to have greater
111 success in generating a widespread interest in the broader scientific community. However, these
112 projects pose several challenges for the diverse science teams. The interests of the various groups
113 involved need to be aligned; participants may lack the required knowledge of other disciplines;
114 traditions and common practices may differ widely between disciplines. Finally, larger teams increase
115 the challenge to communicate and coordinate efforts effectively. The various goals of individual teams
116 call for compromises on several levels, such as drill site selection, subsampling strategies, and choice
117 of analyses (see section 2.1.1.). Life-scientists are typically not familiar with drilling operations and
118 often lack basic geological knowledge whereas earth-scientists may not be acquainted with
119 biochemical or biological procedures. More practically, the difficulty arises that life-scientists do not
120 know exactly how to retrieve the archives they hope to study, and that earth-scientists cannot evaluate
121 applicability and performance of biological methods. Similar problems persist on smaller scales, and
122 given the rapid advancement of many of the individual fields, specialists may even struggle with

123 methodological innovations in their field over the often year-long duration of deep drilling projects,
124 involving the planning, the actual drilling campaign, and the interpretation of the final datasets. These
125 issues are also relevant for core storage, which may affect geological and biological properties
126 differently. Sedimentologists are typically acquainted with long-term changes in sediments after core
127 retrieval, but others may draw erroneous conclusions when linking biological and geological data
128 without accounting for potential contamination, drilling artifacts, decay processes, and other
129 complications (see section 2.1.2.). In general, greater logistic, communicative, and administrative
130 efforts are required with increasing complexity of interdisciplinary projects, and drilling methods may
131 have to be optimized to guarantee the required data quality.

132 Perhaps the most challenging task, however, is to integrate the diverse datasets various teams
133 collect from drilling cores. These datasets typically have differences in resolution, data quality, and
134 dating uncertainty, but combining them is required to answer interdisciplinary questions. Whereas the
135 physical linkage of information directly obtained from sediment cores is, in most cases, relatively
136 straightforward due to the chronological constraints on the data, the challenge grows when primary
137 data, i.e., data generated from sediment cores or in boreholes, are to be linked with external
138 (secondary) data, i.e., data obtained independently of the drilling operation. Examples of secondary
139 data sources include stable isotope information from fossils found in outcrops (see section 2.1.6.) or
140 genetic information from extant species (see section 3.3.).

141 Here we review the types of geological and biological data that can be obtained from ancient lake
142 drilling projects (section 2.) and the methods that can be used to analyze these data against the
143 backdrop of the abovementioned practical and analytical challenges (section 3.). Acknowledging the
144 increasing number of approaches and analyses that can be applied to drilling data, we narrow our
145 focus on data and methods that have a high potential towards integrating geological and biological
146 data and for hypothesis-testing related to interdisciplinary questions. We also provide a retrospect on
147 how the actual drilling operation and conditions of sediment-core storage can affect data and
148 subsequent multi- and interdisciplinary analyses. Although this review focusses on extant ancient
149 lakes, some of the information given might also be applicable to lakes from the past and even young
150 lakes.

151 Our aim is to provide scientists from various disciplines with a background to strengthen
152 interdisciplinary approaches to ancient lake drilling projects. We thus explain data acquisition and
153 analyses in broad terms and provide information as to the underlying fundamental principles that may
154 be equally useful for earth and life scientists. Given this scope, we refrain from detailed discussions
155 that are constrained to a specific field, nor do we provide a historic overview of drilling operations for
156 which other reviews exist (Cohen, 2012).

157 As such, this review intends to encourage scientists from diverse disciplines to join scientific deep
158 drilling projects, and to utilize these unique records of global change during the earth's history for
159 understanding current and future changes on a planetary scale.

160 **2. Data and methods**

161 *2.1. Geological data and methods*

162 *2.1.1. Site selection and drilling strategies*

163 Careful consideration of the drill site(s) and the drilling strategy are a prerequisite to optimize the
164 chances that the goals of a deep drilling project can be reached. Scientific objectives are the foremost
165 criteria for the selection of drill sites and strategies, but financial and time constraints also have an
166 distinct impact. The extensive infrastructure needed and the shipping of highly specialized gear are
167 important cost factors of deep drilling (Fig. 2).

168



169

170 **Fig. 2.** Deep Lake Drilling System of DOSECC (USA) in operation at Lake Ohrid. In this case, the equipment
171 had to be shipped from Salt Lake City (USA) to this inland lake on the Balkans (photo credits: T. Wilke).

172

173 Given a certain budget, the costs for logistics, including the transport of the drilling equipment to and
174 from the lake and daily operational expenses (e.g., labor costs, fuel consumption, daily shuttle of drill
175 team and cores), define the number of drilling days. Technical failures, weather delays, or lithologies that
176 are challenging to drill are difficult to predict and may further reduce available drilling time. In contrast,
177 unexpected drilling progress may provide the opportunity to drill deeper than anticipated or to add
178 additional sites, addressing for example 2nd or 3rd order scientific objectives. For these reasons, a well-
179 calculated budget, which includes sufficient contingency to cover unforeseen costs and/or delays, may
180 significantly increase the chances for success. Finally, site selection may also be influenced by the time
181 required for daily commutes of the drill teams to and from the drill site. Experience has shown that the
182 one-way commute time should not exceed 1.5 h when working in 12 h shifts.

183

184 Most of the scientific criteria for site selection and drilling strategies are reviewed in Cohen (2012).
185 Catchment characteristics, such as the location of inlets, providing terrigenous, clastic, and organic
186 sediments to the basin, the bathymetry and surrounding topography, and the presence of undisturbed
187 successions may affect site selection. However, bathymetry, catchment topography, and the location of
188 inlets may have changed over time, especially in tectonic basins. Most commonly, the site selection is
therefore based on data from reflection-seismic surveys, providing information about bathymetry and the

189 three-dimensional sub-bottom sediment structure, as well as the stability of sediments in target areas. For
190 most scientific questions, areas with undisturbed sedimentation are preferred, but for others, such as the
191 reconstruction of earthquake history, a focus on disturbed successions may be beneficial (e.g., Takemura
192 et al., 2013). Moreover, most drilling campaigns focus on the depocenter of a lake. If biological
193 questions are of interest, such a drilling strategy will almost certainly omit valuable information of the
194 nearshore benthic communities where most of the biodiversity is concentrated (see also section 2.2.2).
195 However, during initial lake phases or massive drops in lake level, shallow water species may have lived
196 nearby the depocenter and thus be preserved in the old sediments just overlying the basement rock (e.g.,
197 Wagner et al., 2014a). Coring in shallower or littoral regions is facing other problems, such as obtaining
198 discontinuous or incomplete records of biotic evolution, hydrological characteristics, and catchment
199 processes. In these environments, sediments from the initial lake phase or low-level periods will be
200 missing.

201 The drilling strategy also includes criteria, such as the number of drill sites, their priorities, or the
202 number of holes per site needed for a composite core. As a general rule, the site with the highest
203 scientific priority should be drilled first, as unforeseeable technical or weather issues may delay the
204 coring progress. However, it may be advisable to start with a site of secondary priority, for example, if
205 the sedimentary characteristics are poorly known or if the main drill site requires high logistic efforts. In
206 many campaigns, the initially retrieved sediment cores have a poorer quality compared to cores drilled at
207 later stages of the operation, because drill and science teams need to develop communication and drilling
208 strategies (e.g., the choice of drill tools) suitable for the encountered lithologies. As core recovery in a
209 single borehole can be as low as 10% (see also section 2.1.2.), parallel coring is needed to fill the gaps.
210 Sometimes 3–4 holes are required to obtain 100% or close to 100% recovery at one site (resulting into a
211 so-called composite core).

212 In some cases, the drilling strategy may need to be adapted to effectively address specific scientific
213 questions or deviating lithological intervals. For example, in studies focusing on subsurface microbial
214 activity, contamination by drilling fluids needs to be minimized and samples may need to be taken
215 directly upon recovery of the cores (see section 2.2.3.). If the exact depth of the recovered sediments is
216 not of highest priority, the hole of the first core drilled for subsurface biosphere studies can be used for

217 borehole logging (section 2.1.3.) and to obtain initial information about lithological characteristics at
218 depth. This facilitates adaptation of drilling strategies for the subsequent holes and sites to be drilled.

219 In summary, several criteria affect site selection and drilling strategies. Because of the specifics of
220 each lake, these criteria need to be balanced carefully based on the exact setting encountered. Meticulous
221 preparation, including the collection of site-specific information, such as data on sediment structure from
222 seismic surveys and lithological characteristics from pilot coring surveys may serve to construct a
223 drilling strategy that can then be evaluated against scientific objectives, budget, logistic requirements,
224 and the available time.

225 *2.1.2. Coring techniques and data recording*

226 The main lake coring tool is a piston-type corer, which allows penetration under hydraulic power into
227 generally cohesive sediments of various aggregation up to a pre-determined length. This tool permits
228 collection of the least 'contaminated' or 'disturbed' type of core as it is pushed into the formation ahead
229 of any disturbance caused by the rotational drill bit and sediments are collected before they get into
230 contact with drilling fluids. Before the actual drilling, the piston corer is sealed with a piston to avoid
231 contamination of the drill string with surface lake water and cuttings or cavings, which may be in the
232 borehole before and during the core collection, to enter the coring tube.

233 In soft to firm clays, sample recovery can be 100%. The slight under-pressure caused by the piston
234 avoids shearing of sediment packages in front of the corer during penetration, but it leads to expansion of
235 the core, particularly at its top and may promote degassing. In more granular sediments, compression
236 rather than expansion is common due to grain packing. Moreover, as the drill string is frequently
237 advanced to the next firing point by the length of recovery rather than the length of the core barrel, there
238 can be overlap of 'strata' between two core runs, with the top of a subsequent sample consisting of
239 cavings from the previous sample. It is not easy to completely clean the borehole in this type of sediment
240 and the coarser and less cohesive it is, the more likely it will be that full recovery of the section will not
241 fill the core barrel due to packing and dewatering occurring. Again, recalibration of any subsample depth
242 with regard to the composite borehole log and marker horizons will ensure that correct depths for
243 correlation with other scientific data results are achieved. Although the core length may not reflect the

244 full sampled length, the stratigraphic coverage typically will be complete and essentially free from other
245 types of mechanical disturbance and fluid mixing. Therefore, piston cores allow for the best stratigraphic
246 resolution of all drill-sampling method.

247 Other frequently used lake drilling tools are the extended nose corer and the alien corer. Both of those
248 tools require a rotation while coring. The core is cut and collected into the core barrel in close proximity
249 to the drill bit while it is operating and, hence, while fluid discharge is required to keep the drill bit cool
250 and clear of cuttings. This makes it difficult to allow the core an undisturbed passage into the core barrel,
251 and results in artifacts by the drilling and the use of drilling fluid. These artifacts depend also on the
252 lithology of drilled sediments, but there may be mixing of drilling and formation fluids, or selected
253 portions of the core may be washed away, and disturbance to the structure and stratigraphy of the core
254 may also occur.

255 The extended nose corer is used when the material in the stratigraphic section is non-cohesive or
256 friable but becomes too coarse or compacted to be able to operate a piston corer for any reasonable
257 length into the subsurface. To use this corer, the sediment needs to be still soft or granular enough to
258 allow extending a thin diamond bit nose ahead of the main bit and main fluid flush when progressing the
259 hole. Ideally, the core will enter the extended bit section without too much interference or influence from
260 the rotation and drilling fluid. If there is any movement of the drill string due to heave of the platform
261 while drilling, then discs of core and cuttings may also occur and show up as bands in the core. Recovery
262 is unlikely to reach 100% even under ideal conditions.

263 The alien or rotary core barrel is essentially a method of collecting cores from hard, cohesive
264 sediments. It may selectively wash away material that is too soft or non-cohesive before it can be
265 properly cut to enter the core barrel. It may also induce drilling fractures in otherwise good quality rock.
266 There is always an interaction between the drilling and formation fluids as the core is being cut and
267 collected. Recovery will range from poor (<10%) to excellent (100%), depending on lithology, weather
268 conditions, and type of drill bit. However, good recovery does not equate to quality core for laboratory
269 analyses given the earlier highlighted artifacts. The length of the recovered core may not be a good
270 reflection of the length of the stratigraphic section cored. Harder material is most likely to be recovered
271 while soft material can be crushed or washed away, thus reducing its occurrence in the recovered

272 stratigraphy. In these cases, composite borehole log interpretation, utilizing borehole geophysical and
273 Multisensor Core Logger (MSCL) data, is important to indicate what sections are recovered and how
274 they are to be placed in the actual stratigraphy. Stratigraphic resolution with this type of corer will be
275 very good in hard materials but will quickly diminish where uncemented, friable, or non-cohesive
276 sediments occur in the section. Subsamples of such cores, hence, need to be carefully extracted.

277 In general, core portions used for bulk analyses should be disassembled and cleaned of foreign
278 materials before use. If, in later years, bulk samples are made available for analyses by scientists that
279 were no part of the drilling team, it is important that the core's full history is known, so that anomalous
280 or unexpected results can be interpreted in the light of possible artifacts of the drilling, storage or
281 preparation of the material. Detailed drilling and core interval logs should be held with the composite
282 geological core log and identify, as a minimum, the type of drilling and coring tools used, and details of
283 any drilling mud/fluid utilized. The latter information is important to eliminate geochemical anomalies in
284 scientific results. For example, contact of long-deposited sediments with current lake water may leave a
285 signature in the stored core. Although the drilling mud will not directly affect fossil material (sections
286 2.2.1., 2.2.2.), unless there are secondary reactions during storage, it may affect microbial activity,
287 metabolites, and decay products as some types of drilling mud contain guar, which is a food source. For
288 microbiological studies (see section 2.2.3.), tracers used to detect 'contamination' need to be
289 documented. Moreover, if microspheres were utilized, they may still be present in some samples.

290 Subsamples of sediment cores from scientific drilling projects are increasingly being used for
291 multidisciplinary studies in a much wider scientific aspect than that of the original project and thus
292 particular emphasis should be placed on understanding the circumstances in which the original data were
293 collected and records were archived (for a review see Cohen, 2012). Any subsample from a core section
294 needs to have the depth of the core section as an unmistakable criterion, as correlation depths derived
295 from the correlation of cores from parallel holes (so-called core composites) may change with results
296 from ongoing measurements (e.g., high resolution logging). Ideally, these core composites are
297 recalibrated to borehole logging data prepared from a number of boreholes at the same site. However,
298 significant marker horizons (e.g., tephra deposits), being used as the correlation basis for the composite
299 borehole log are not always available.

300 Storage in controlled and cool temperatures will minimize moisture loss, bacterial activity (see
301 section 2.2.3.) and ancient DNA decay (see section 2.2.6.), and allow physical property measurements to
302 be extended by a few months. Cooler temperatures will also slow down any secondary chemical
303 reactions or existing core alteration.

304 Drilling operation and coring methodology define initial core quality and determine the degree of
305 physical or chemical 'contamination' that may be anticipated. Archiving and storage imprints, further
306 characteristics on the core, and accurate subsampling interpretation require a full tracking record of the
307 core from collection to interpretation with subsample positions clearly archived in relation to the final
308 composite borehole log. These logs should also show where subsamples have been taken from and
309 regularly updated as new information is generated.

310 2.1.3. Borehole logging

311 Borehole logging is the process of measuring physical, chemical, and structural properties of
312 penetrated geological formations via tools that are lowered into a borehole on a wireline cable. It
313 provides *in situ* information about the physical properties of the rock or sediment strata and groundwater.
314 Borehole logs deliver a continuous record that provides information on the lithological changes with a
315 precision of decimeters to a few centimeters. Since it allows depicting actual depth and petrophysical
316 characteristics, information from borehole logging is often used in combination with seismic reflection
317 data to construct geological models. In addition, the combination of downhole logging data and
318 petrophysical datasets from several drill holes and cores from the same site are essential to construct a
319 composite lithological log. More recent applications include the derivation of paleoclimatic indicators
320 and cyclostratigraphic analyses.

321 The main components of logging equipment are a surface unit, winch, cable, and logging tools
322 equipped with variable detectors and/or sensors. The surface unit is used to control the measurements,
323 including the movement of the tool in the borehole. It also provides the energy supply to the tool and
324 records, displays, and stores the data generated in the borehole. The depth of the measurement in the
325 borehole is independently recorded by a gauge on the winch. To allow depth correlation between all
326 logging runs, each tool is equipped with a gamma ray sensor, which records the formation's natural

327 gamma radioactivity caused by its occurring contents of uranium (U), thorium (Th), and potassium (K).

328 A number of tools have been developed over the recent decades that maximize the number of
329 physical parameters that can be measured in slim boreholes. The equipment, field application, and
330 analytical methods have been described by, e.g., Ellis and Singer (2007) and Rider and Kennedy (2011).
331 The most important tools/physical aspects of borehole logging in a lake drilling project are spectral
332 gamma ray (natural gamma ray plus contents of U, Th, and K), magnetic susceptibility, resistivity,
333 acoustic velocity, vertical seismic profiling (VSP), dipmeter, and caliper (borehole diameter and
334 orientation). Furthermore, tools that register the density, neutron porosity, and the content of a selection
335 of geochemical elements of the drilled formations are available. These tools emit ionizing radiation and
336 contain either radioactive sources or neutron accelerators. The regulations for the import and export of
337 these tools are complex and differ from country to country, which regularly limits their use in ancient
338 lake drilling projects. Tools based on optical methods like video cameras or optical televiewers exist as
339 well, but they cannot usually be operated in lake drillings because the drilling fluids are not translucent.

340 The various downhole logging methods together with a sensor configuration adapted to the
341 measurement conditions is used to obtain data from a limited, irregular rock volume. The vertical and
342 radial extent of this volume is influenced by the borehole diameter, the physical properties of the content
343 of the borehole, the ratio of the borehole diameter to the diameter of the tool, the position of the tool in
344 the borehole, and the design of the tool (detector size, electrode spacing, transmitter-receiver spacing,
345 radioactive source-detector spacing). Thus, each tool has a characteristic depth resolution and an average
346 radial depth of investigation under the given conditions.

347 The logistical effort involved with transport and installing logging equipment at the drill site can be
348 high, particularly when the gear has to be transported to a floating drilling barge. These barges are
349 typically not equipped with heave compensation, so that logging operations need to be conducted during
350 appropriate weather conditions. Final decisions about holes to be logged during or at the end of drilling
351 operations can change quickly and are dependent on the overall progress of the drilling operation.

352 Downhole logging measurements in ancient lake drilling projects are typically made in
353 unconsolidated sediments. Due to the specific sensor requirements, most physical parameters have to be
354 measured in an open hole. To reduce the impact of potential borehole wall collapses, logging is

355 performed in borehole sections. Their number and individual length (down to 30 m) need to be defined
356 in close cooperation with the drilling supervisor and leads to a significant prolongation of the logging
357 time. The logging speed depends on the tools/sensors used, and has a large influence on the quality of the
358 data. Typical logging speeds are between 60 m h⁻¹ and 600 m h⁻¹. The sampling interval also influences
359 data quality. It is typically 5–10 cm, with a vertical resolution of about 20 cm.

360 The quality of the borehole logs are dependent on the measurement conditions and the depth
361 resolution of the tool. The conditions of logging are determined primarily by the borehole geometry, the
362 physical properties of the drilling fluid (density, electrical resistivity, pH, neutron braking, and
363 absorption properties), and the properties and size of the flushed and invaded zone. Borehole logging is
364 thus an important and versatile tool in ancient lake drilling projects and applications include:

365 i) Lithological classification of the strata penetrated by a borehole.

366 Characteristic physical rock parameters, especially when combined, can be used to determine or
367 confirm the lithology of the rocks or sediment successions. The most useful parameters for this are
368 gamma radiation, magnetic susceptibility, resistivity, acoustic velocity, and caliper.

369 ii) Site-to-site/hole-to-hole correlation of lithological units.

370 Correlation of lithological units are crucial for obtaining core-composites and for linking stratigraphic
371 positions across boreholes (see section 2.1.2.). Best suitable are gamma-radiation and magnetic
372 susceptibility. Furthermore, magnetic susceptibility data has the potential for identifying tephra layers
373 that are important chronostratigraphic marker horizons in sediment successions.

374 iii) Paleoenvironmental/paleoclimatic reconstructions.

375 Paleoclimatic indicators were derived from physical properties of the sediments from lakes
376 El'gygytgyn (Nowaczyk et al., 2013), Van (Baumgarten and Wonik, 2014), and Ohrid (Baumgarten
377 et al., 2015). Sediment records from these lakes display strong shifts in physical and chemical
378 properties with depth that are tied to different environmental/climatic states.

379 iv) Cyclostratigraphic analyses.

380 Contrasting physical properties and therefore changes in the sediment characteristics can trigger
381 cyclic changes in the logging data. Cyclostratigraphic analyses in lakes based on physical parameters
382 from cores and/or boreholes were conducted by Nowaczyk et al. (2013), Baumgarten and Wonik

383 (2014), and Baumgarten et al. (2015). Under favorable sedimentation conditions, results from
384 cyclostratigraphic analyses based on downhole logging data even allow estimates of varying
385 sedimentation rates and establishment of reliable age-depth relationships (see section 2.1.4.).

386 v) Time-depth conversions.

387 Measurements of acoustic velocities in boreholes are useful for the interpretation of shipborne
388 seismic reflection surveys. Therefore, for the geological modeling based on seismic data, downhole
389 logs (acoustic velocity and VSP) can provide essential information necessary to develop time-depth
390 conversions for seismic datasets.

391 vi) Inference of structural/tectonic features.

392 Important information needed for tectonic and structural geology investigations include the dip angle,
393 strike direction, and angle of formations. These parameters can be inferred in the borehole using a
394 dipmeter tool. Layers with thicknesses as small as 1–2 cm can be detected.

395 vii) Hydrogeological and geotechnical parameters (e.g., clay content, porosity).

396 Data registered from the acoustic velocity logging tool can be used to determine porosity.

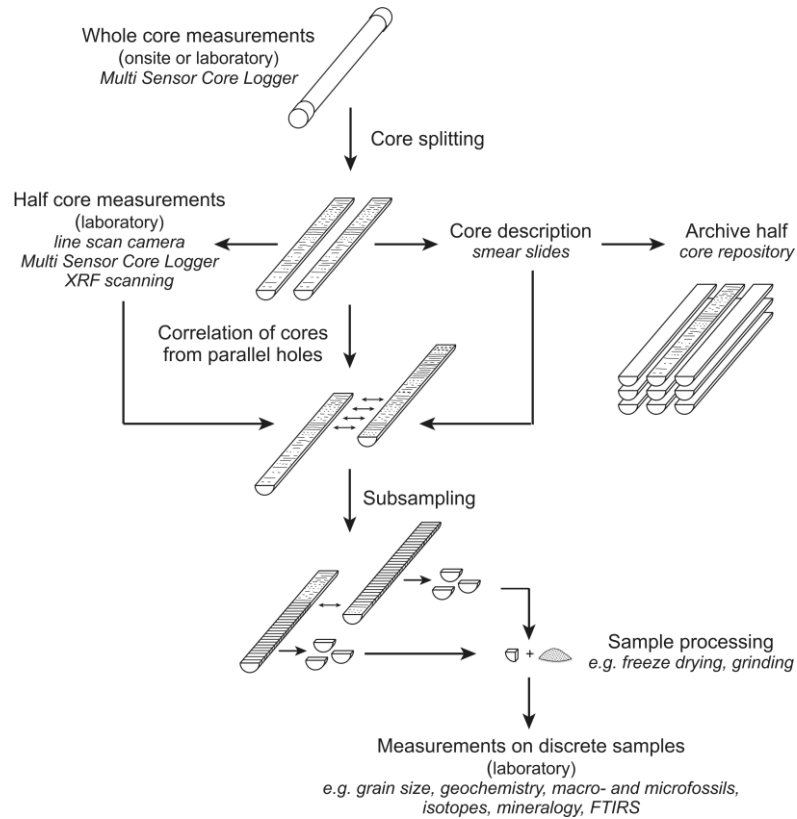
397 Assumptions about the lithology and fluid properties based on local knowledge or other
398 measurements have to be made to estimate porosity. Spectral gamma ray logs may help to distinguish
399 between different clayey sediments based on their U, Th, and K content. Acoustic velocity logs via
400 the compression modulus provide indirect information about the storage coefficient of the rock or
401 sediment successions.

402 Due to technical difficulties, micro-resistivity imagers and nuclear magnetic resonance (NMR)
403 equipment have not been applied in slim boreholes so far (for details of the methods see, e.g., Kenyon,
404 1997; Coates et al., 1999; Dunn et al., 2002). With the former method, detailed, high-resolution spatial
405 data providing structural and textural information can be obtained from the measurement of resistivity
406 with a large number of electrodes (Lovell and Parkinson, 2002). Using a NMR tool, a real time
407 permeability log may be possible. Also, the application of interpretative methods for hydrocarbon
408 exploration and basement rock data (e.g., multivariate statistics) to ancient lake drilling projects can
409 provide further possibilities to combine geophysical data from boreholes with other geophysical and
410 geochemical datasets towards more holistic analyses in the future.

411 2.1.4. *Sediment-core data*

412 Standard measurements on sediment cores can be separated into those conducted on whole core
 413 sections, on split core surfaces, and on discrete samples (Fig. 3).

414



415

416

417 **Fig. 3.** Generalized scheme of core processing for basic sedimentological analyses. Individual steps and analytical
 418 methods may vary across projects.

419

420 Whole core measurements are typically performed just after drilling on site and/or shortly before core
 421 opening in the laboratory. For these purposes, so-called multi-sensor core loggers (MSCL), which can be
 422 described as a logging bench unit, are commonly used. MSCL systems allow stepwise measurements of
 423 geophysical parameters at centimeter resolution and are commonly equipped with variable
 424 sensors/detection systems for magnetic susceptibility, gamma ray density (GRAPE), P-wave velocity,
 425 natural gamma spectroscopy, and resistivity (e.g., Weber et al., 1997). These datasets provide a first,
 426 rough overview about changes in sediment composition (density, detrital/magnetic mineral content, U-,
 427 Th-, K-concentrations) before the cores are opened and sediments are exposed to the atmosphere. They

428 are useful to establish hole-to-hole correlations, to identify gaps in the core recovery, and to assist
429 adjusting the drilling strategy. P-wave velocity data are needed to establish time/depth conversions for
430 acoustic and seismic reflection datasets, thus allowing for more sophisticated interpretations of seismic
431 units and their correspondence to changes in lithology. Furthermore, obtaining these whole core
432 measurements is a prerequisite for geomicrobiological sampling (see also section 2.2.4.), which needs to
433 be undertaken in the field shortly after core recovery and consumes entire core sections. Hence, if
434 geomicrobiological datasets need to be tied to a stratigraphy/chronology or compared to other analyses
435 obtained on overlapping intervals later on, this can only be achieved using whole core MSCL datasets.
436 MSCL datasets can be biased by cracks in the sediments due to gas expansion or by incompletely filled
437 liners, such as is common when using rotational drilling tools (see section 2.1.2.).

438 Macroscopic lithological core descriptions on split core surfaces, supported by microscopic smear
439 slide analyses of major biogenic and minerogenic sediment components, are used to define major
440 lithologic (i.e., pelagic sediments, mass wasting deposits, tephra layers) and sedimentary/stratigraphic
441 units. This information is not only important to determine changes in the depositional environment but
442 also allows more targeted lithotype-specific subsampling strategies to be developed. Detailed and
443 spatially highly resolved (down to 0.2 mm) geochemical analyses on split core surfaces can be
444 performed using X-Ray Fluorescence (XRF) core-scanners. Newer systems allow the analysis of major
445 and minor elements heavier than Al at high precision and permit, besides inference of relative changes in
446 concentration, also quantification upon calibration (e.g., Russell et al., 2014). Relative changes in
447 intensities as well as ratios of specific elements are useful indicators for changes in sediment
448 composition (e.g., Melles et al., 2012), sediment provenance (e.g., Kylander et al., 2013), redox
449 conditions (e.g., Naeher et al., 2013; Costa et al., 2015), diagenetic overprinting, grainsize/depositional
450 processes (e.g., Bloemsma et al., 2012; Francke et al., 2016), and to identify the occurrence of tephra or
451 cryptotephra layers (e.g., Vogel et al., 2010a). The quality of the data depends on the analytical time and
452 on several sedimentological characteristics, such as surface smoothness, water content, or grain size
453 distribution. XRF scanning of marine and lacustrine core sequences has become a standard method to
454 obtain a rapid, spatially highly resolved overview of changes in element composition, but the method is
455 time consuming at high resolutions. Requiring an analysis time of up to several hours per meter sediment

456 core, the scanning of several sediment sequences from one site with a total length of hundreds of meters
457 may take several weeks. Moreover, due to the high scanning resolution, the amount of data produced is
458 immense and often requires smoothing or filtering in order to facilitate the identification of major trends
459 of element variation in sediment cores and to reduce background noise. Therefore, careful selection of
460 the resolution based on other data, e.g., reduction of resolution if a section of sediments appears
461 homogeneous, for example, due to bioturbation or mass movement processes, can significantly reduce
462 measurement time and facilitate data handling.

463 Analytical methods on discrete samples comprise a broad range of granulometric, geochemical, and
464 biological methods. Much variation exists in the amount of preparation and time that are required to
465 apply individual methods, e.g., sample preparation or identification of micro- and macrofossils can be
466 very time consuming. Significant differences also exist in the required amount of sediment needed to
467 perform each analysis. Some geochemical methods require only a few milligrams of powdered material,
468 but others, such as paleontological studies may need several grams of sediment or more. Both labor-
469 intensity and availability of material affect the sampling resolution, with the separation between
470 subsequent samples usually varying from one centimeter to several decimeters.

471 Granulometric analyses are used to obtain information on sediment transport history and energy
472 (fluvial, aeolian, pelagic, gravity driven). The measurement of grain-size distributions can be done
473 relatively fast with a laser particle size analyzer for sand- to clay-sized fractions, as are common in
474 lacustrine sediments. However, sample preparation is necessary to extract information on transport and
475 depositional processes unbiased by autochthonously derived sediment components. Sample preparation
476 can be labor-intensive if removal of organic matter, carbonates (authigenic, biogenic), biogenic silica
477 from diatoms, phytoliths, or sponge spicules is required. Standard geochemical analyses often comprise
478 the measurement of total organic carbon (TOC), total inorganic carbon (TIC), total nitrogen (TN), total
479 sulphur (TS), and require only relatively low amounts of powdered material. TOC and TN
480 concentrations are useful indicators that provide estimates on changes in intralacustrine productivity
481 (e.g., Wagner et al., 2009; Vogel et al., 2010b) and/or changes in the supply of terrestrial organic matter
482 (e.g., Meyers, 2003). However, TOC and TN concentrations in the sediment are not only influenced by
483 their initial fluxes but also by the degree of post-burial remineralization, which in most settings is

484 controlled by lake-mixis and the availability of oxygen in the hypolimnion and top sediments (e.g.,
485 Melles et al., 2007). TIC concentrations are a measure of the amount of carbonate, which can be present
486 in biogenic, endogenic, detrital, and/or authigenic form in lacustrine sediments. Calcite and aragonite
487 (both CaCO_3) are the predominant carbonate phases in most freshwater settings and changes in their
488 concentrations are usually driven by temperature, productivity, and hydrological variations (e.g., Kelts
489 and Talbot, 1990; Wick et al., 2003; Wagner et al., 2009; Vogel et al., 2010b). However, other carbonate
490 phases, such as siderite (FeCO_3), dolomite ($\text{CaMg}(\text{CO}_3)_2$), and/or ankerite ($\text{Ca}(\text{MgFeMn})(\text{CO}_3)_2$), can
491 also be present in certain settings (e.g., De Decker and Last, 1988; Stevens et al., 2012; Lacey et al.,
492 2015). Moreover, the measurement of isotopes, such as carbon or oxygen isotopes (see section 2.1.6.),
493 require relatively small quantities of sediment if the concentration of biomineralized carbon or oxygen is
494 sufficient.

495 A new analytical method that has been applied in recent deep drilling projects for the analysis of
496 discrete samples is Fourier Transformed Infrared Spectroscopy (FTIRS). FTIRS is a relatively fast and
497 cost efficient method, which requires small sample amounts and can be used to infer absolute
498 concentrations of biogenic silica (bSi), carbonate (TIC), and organic matter (TOC) with a single
499 measurement (Vogel et al., 2008; Meyer-Jacob et al., 2014a). The low processing time might be
500 important for the continuous analysis of long drilling sequences (e.g., Meyer-Jacob et al., 2014b) and for
501 high-resolution studies of specific time slices in these sediment records (e.g., Cunningham et al., 2013).
502 Further applications of FTIRS involve the determination of relative changes in the abundance of
503 different carbonate phases (Lacey et al., 2016).

504 Overall, the combination of information from individual proxies and future progress in the
505 development of analytical methods as well as scanning and logging techniques will significantly help to
506 improve the study of environmental changes from sedimentary records in ancient lakes.

507 *2.1.5. Age-depth models*

508 Creating a reliable and robust chronological framework is fundamental for drilling studies across
509 disciplines and, hence, also for the synthesized interpretation of paleoenvironmental, climatological, and
510 biological data. Age-depth relationships in non-marine records are commonly established by combining

511 absolute chronological information from radiometric and magnetic dating methods, and from varve
 512 counting with chronostratigraphic information derived from comparisons of a proxy response to a
 513 reference record (e.g., Nowaczyk et al., 2013; Stockhecke et al., 2014; Francke et al., 2016). Similar to
 514 dendrochronology (tree ring counting), varve counting provides robust age-depth control points as
 515 varves consist of thin (millimeter scale), characteristic seasonal summer and winter deposits (so-called
 516 laminae). However, the preservation of varves in lacustrine sediments may depend on several factors,
 517 such as the absence of sediment-dwelling organisms or the presence of anoxic bottom water conditions.

518 Obtaining absolute ages from radiometric dating techniques (Table 1) or techniques that utilize
 519 radiometrically induced lattice effects in certain mineral phases requires the presence of suitable/datable
 520 materials in the studied sediment sequence (reviewed in Bradley, 2014). In addition, the different dating
 521 methods cover different time ranges. Consequently, various radiometric dating techniques are typically
 522 used in combination and choices depend on the age range covered in a core, the anticipated/estimated
 523 age of a specific dating point, and the presence of suitable materials. Whereas the relative error of some
 524 of these techniques is small, absolute uncertainties increase in the deeper sediment record and obtained
 525 dates can be biased by a variety of physical and chemical effects. Nevertheless, dates from radiometric
 526 methods provide the most precise chronological tie points, and are therefore introduced as 1st order
 527 constraints in age-depth calculations.

528

529 **Table 1**

530 Selected absolute dating techniques for sediment cores. ESR = Electron Spin Resonance; OSL = Optically
 531 Stimulated Luminescence. ¹⁰Be refers to burial dating.

532

	¹⁴ C	Ar/Ar	U/Th	U/Pb	OSL	ESR	¹⁰ Be
Material	Organic matter/ carbonate fossils	Volcanic glass	Endogenic/ authigenic carbonates	Endogenic/ authigenic carbonates	Quartz/ feldspar	Carbonates	Quartz
Age range (My)	0–0.05	~0.01–4,600	0–0.5	1–4,600	0.001–0.2	0.001–0.2	~0.01–4
Precision (%)	1–5	5	1	0.1–1	~10	~10	10–20

533

534 Additional 1st order constraints can be obtained by correlating the chemical composition of tephra

535 layers to the volcanic eruption of which the age is known or to tephra layers with identical composition
536 found in other, well-dated sediment sequences (e.g., Wulf et al., 2008; Sulpizio et al., 2010; Vogel et al.,
537 2010a; Leicher et al., 2015). However, an unequivocal correlation of tephra layers to known eruptions
538 based on their chemical composition is, due to the often encountered similarity in glass shard chemistry
539 of ejectas from different eruptions, not always straightforward. In recent studies, these limitations have
540 been partly overcome by additional measurement of the trace element and isotope chemistry of glass
541 shards (e.g., Sulpizio et al., 2013; Insinga et al., 2014; Albert et al., 2015; Tomlinson et al., 2015).

542 Further independent age control can be provided by paleomagnetic data. Polarity reversals and
543 excursions of the earth's magnetic field are common in the geological record (Merrill et al., 2006).
544 Moreover, the intensity and vector (secular variation) of the Earth's Magnetic Field varies within
545 magnetic chrons. As magnetic particles will be aligned during and after settling on the lake floor, their
546 orientation in the sediment core depends on the direction and strength of the ambient magnetic field
547 during or shortly after deposition (depositional remnant magnetization, DRM). Alignment of particles
548 during deposition is, however, not always an instantaneous process and different factors, such as the
549 lock-in depth of DRM (i.e., the sediment depth where magnetic particles are aligned to the ambient
550 magnetic field, important for sedimentation rates of $<10 \text{ cm ky}^{-1}$; Roberts and Winklhofer, 2004) and
551 early diagenetic formation of ferrimagnetic minerals (important in low oxygen environments) should be
552 considered (see Tauxe, 1993; Roberts et al., 2013 for more details). Age information from well-dated,
553 paleomagnetic reference records can be transferred to the analyzed sediment sequence, if similar
554 magnetic features were identified (Peck et al., 1996; Frank et al., 2002; Nowaczyk et al., 2013). A
555 number of magnetic reversals and excursions have been dated using absolute radiometric dating methods
556 (Laj and Channel, 2009 and references therein). This is, however, not yet the case for the entirety of
557 these events in the geological history. In some cases, the chronology of the paleomagnetic reference
558 records is based on synchronization of marine benthic oxygen isotope data of either individual or stacked
559 records with orbital parameters, standard oxygen isotope stacks (SPECMAP and LR04), or ice-core
560 records (GISP2). Where absolute age control of paleomagnetic events based on radiometric methods is
561 absent, ages derived from magnetostratigraphy should not be regarded as absolute age control points.

562 Downhole magnetostratigraphic measurements can be performed even with logging-while-drilling

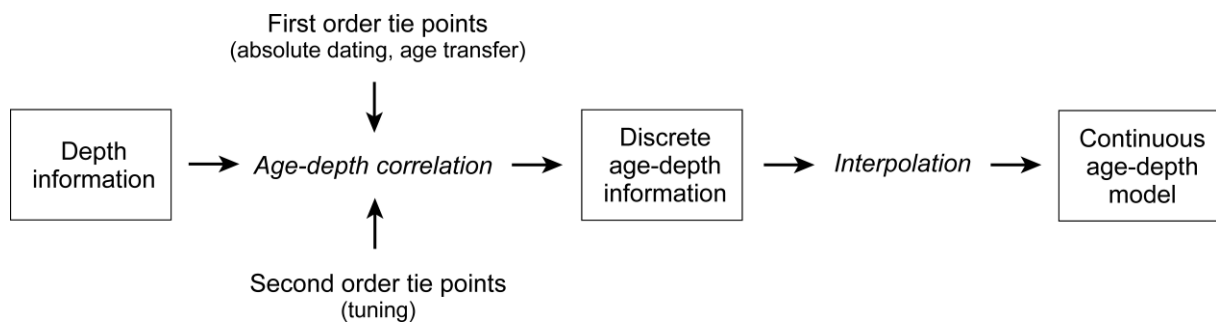
563 tools, such as the geological high-resolution magnetic tool (GHMT), which is based on high-precision
564 measurements of the total magnetic field and the susceptibility field with a magnetometer. Combining
565 these two measurements with the Earth's total magnetic field at the surface close to the borehole allows
566 deriving *in situ* remnant magnetizations of the drilled sedimentary units (Luthi, 2001; Paulissen et al.,
567 2011).

568 Correlation of a proxy response measured in a sediment record to local insolation or to the global
569 benthic foraminifera oxygen isotope stack LR04 (Lisiecki and Raymo, 2005), which primarily displays
570 ice-volume/sea-level changes during the Quaternary, is a chronostratigraphic method that is often used to
571 refine age-models when independent age control points are sparse. At Lake Ohrid, for example,
572 variations in the length of local insolation during summer and winter trigger changes in the primary
573 productivity and mixing in the lake (Francke et al., 2016). These variations affect the TOC content of
574 Lake Ohrid's sediments, which allow tuning the total organic carbon concentrations with the Earth's
575 orbital parameters (i.e., local summer insolation and winter season length). At Lake El'gygytgyn, various
576 stratigraphic parameters, which are related to the redox conditions at the lake floor and to the climatic
577 conditions in the surroundings of the lake, show variation synchronous with the northern hemisphere
578 insolation and the global benthic isotope stack LR04 (Nowaczyk et al., 2013).

579 Tuning of a proxy response against reference records requires a thorough understanding of the
580 processes that shape the proxy response to interpret possible leads and lags compared to the reference
581 dataset (Prokopenko et al., 2006). This is particularly important in lake studies as proxy/forcing
582 relationships are strongly dependent on site characteristics. Ideally, leads and lags between proxy
583 responses and reference records should be tested on horizons for which absolute time markers are
584 available in both the studied and reference records. Zanchetta et al. (2015) showed that tuning against an
585 absolutely-dated reference record from the vicinity of the study site can improve the understanding of a
586 proxy response compared to insolation and ice-volume reference records. Furthermore, this approach
587 may enable the identification of the synchronicity or asynchronicity of climate events compared to the
588 reference records and other, absolutely-dated regional records. The feasibility of such endeavors,
589 however, depends on the availability of suitable reference records in the region.

590 Chronological constraints adopted from tuning against a reference record comprise two potential

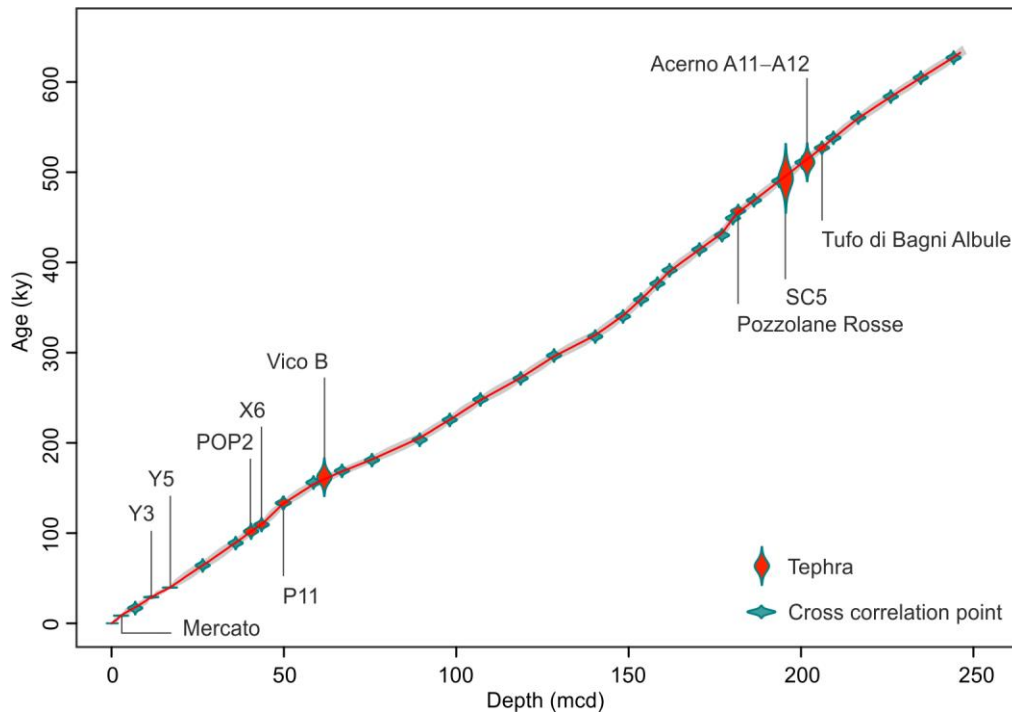
591 errors: Uncertainties introduced by the tuning and the uncertainty of the age model of the reference
 592 record. Therefore, these tie points are not independent and introduced as 2nd order constraints (e.g.,
 593 Nowaczyk et al., 2013; Francke et al., 2016). For example, the chronology of the global benthic isotope
 594 stack LR04, which is frequently used as reference record (see above), comprises uncertainties in a range
 595 of ± 4 –40 ky for a timeframe of 0–5.3 My, as the age model of LR04 is based on tuning the benthic
 596 isotope data to the 21 June insolation at 65°N (Lisiecki and Raymo, 2005). For records younger than 1
 597 My (a relevant range for most lake sediment studies), the error of LR04 is estimated close to ± 4 ky.
 598



599
 600
 601 **Fig. 4.** Flow chart illustrating the required steps to establish an age-depth model.

602
 603 Once a sufficient number of data points have been obtained via the outlined methods, a quantitative
 604 age-depth model can be calculated by interpolation between the individual age control points (Fig. 4).
 605 For this purpose, various interpolation methods, such as linear interpolation, linear regression, polymodal
 606 interpolation, or smooth spline interpolation, can be used. Considering the sedimentological
 607 characteristics, expected sedimentation rates, and other stratigraphic information, including the position
 608 of event layers (mass wasting deposits, tephra layers) and hiatuses forms the basis for deciding which
 609 interpolation method is most appropriate. Linear interpolation implies abrupt changes in the
 610 sedimentation rate at each age control point (cf. Blaauw, 2010; Blaauw and Christen, 2011), which, in
 611 most cases, is not supported by the lithological characteristics and stratigraphic information from the
 612 studied sediment succession. Polymodal or smooth spline interpolation methods calculate more gradual
 613 changes of sedimentation rates and are often more suitable for age-depth modeling. Non-linear changes
 614 in sedimentation rates between two age points are incorporated in new age-depth modeling software, e.g.
 615 Bacon v2.2 (Blaauw and Christen, 2011), which has been applied only recently for age-depth modeling

616 on long lacustrine sediment sequences (e.g., Shanahan et al., 2013; Francke et al., 2016). It uses
 617 Bayesian statistics and Markov Chain Monte Carlo iterations to infer the accumulation history based on
 618 a priori assumptions about the sedimentation rate and its variability over geological time (Fig. 5).
 619 Another advantage of this approach is that uncertainties of tuning and independent age points can
 620 directly be included into the age-depth modeling (Fig. 5).
 621



622
 623 **Fig. 5.** Age model of the DEEP site sediment sequence from Lake Ohrid (Macedonia, Albania) down to 247.8
 624 meter composite depth (mcd) corresponding to an age of 637 ky. Modified from Francke et al. (2016).
 625

626 In summary, age-depth models, which reveal the relationship between increasing sediment depth and
 627 age (Fig. 5), are a prerequisite to integrate and interpret biological and geological data. However, as each
 628 dating technique has its specific limitations, their respective uncertainties should be considered during
 629 age-depth calculations (Blaauw and Heegaard, 2012).

630 2.1.6. Stable isotopes

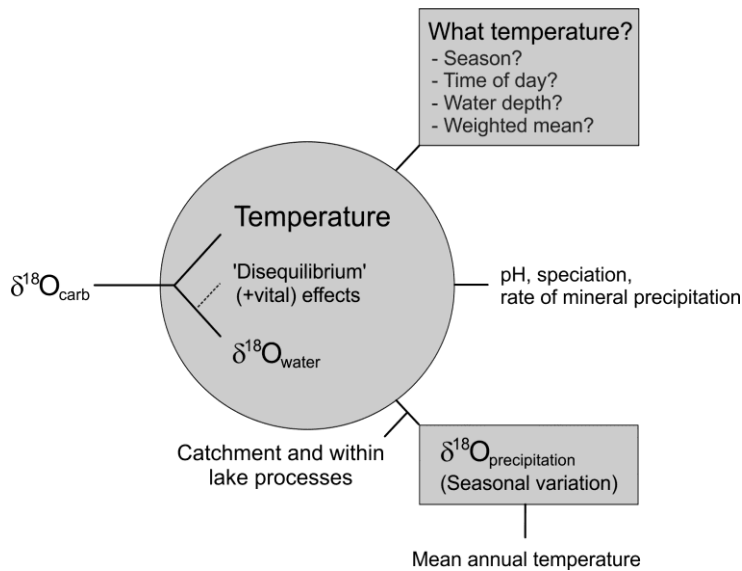
631 The analysis of stable isotope ratios in mineralized components from lake sediments (Leng and
 632 Marshall, 2004), are well established in paleoclimatology, paleolimnology, and limnogeology since the
 633 early work of McCrea (1950) and Urey et al. (1951). Stable isotopes (in particular $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$; Leng

634 and Marshall, 2004) are incorporated into a number of different components that precipitate or grow in
635 lake waters and subsequently get deposited within the sediments, for example shelly materials, fine
636 grained calcium carbonate crystals (a chemical precipitate called endogenic carbonate), and diatoms
637 (siliceous algae, see also section 2.2.2.). Paleoclimate studies commonly use changes in the isotope
638 compositions throughout the sediment succession to infer changes in either temperature ($\delta^{18}\text{O}$) or the
639 isotopic composition of lake water ($\delta^{18}\text{O}/\delta^{13}\text{C}$). With $\delta^{18}\text{O}$, the latter could be a function of changes in/or
640 at the source of water to the lake (changes in moisture source) or the precipitation/evaporation balance.
641 By looking at changes in $\delta^{18}\text{O}$ through time, and depending on the characteristics of the lake in question,
642 it is possible to reconstruct an aspect of paleoclimate for a particular location. With $\delta^{13}\text{C}$, changes are
643 usually ascribed to the source of the dissolved carbon ion or changes in productivity of the aquatic plants
644 and phytoplankton, which preferentially utilize ^{12}C (Leng and Marshall, 2004).

645 In ancient lakes, where the isotope composition of the lake water has been demonstrated to reflect
646 mean annual precipitation, it has been shown that a signal of climate change can be determined from the
647 analysis of biogenic calcite from the shells of benthic ostracods, which live below the thermocline and
648 are thus not affected by seasonal shifts in temperature or evaporation. For endorheic lakes, the isotopic
649 composition of the carbonate has been shown to reflect lake-level, most significantly between glacial-
650 interglacial cycles. In Lake Ohrid, for example, these cycles suggest that the lake has been subject to
651 hydroclimate fluctuations on orbital and millennial timescales (Lacey et al., 2016).

652 Despite an increase in the application of stable isotope techniques in paleolimnology, interpreting
653 stable isotope results can be challenging. These challenges may be general or specific to a core, lake, or
654 basin. The vast majority of studies focus on oxygen isotopes because their interpretation can be linked
655 directly or indirectly to climate change. Factors that have an influence on the oxygen isotope
656 composition of the lake water ($\delta^{18}\text{O}_{\text{lakewater}}$) are given in Fig. 6. The oxygen isotope composition of water
657 in hydrologically open lakes ($\delta^{18}\text{O}_{\text{water}}$) will dominantly reflect the isotopic composition of the
658 precipitation received by the lake ($\delta^{18}\text{O}_{\text{precipitation}}$ or δp). Many studies have shown that the $\delta^{18}\text{O}$ signature
659 of mean annual precipitation varies globally between regions and covariation in δp (and δD) defines a
660 global meteoric water line (Craig, 1961). Outside the tropics, where ‘amount’ effects (relating to intense

661 precipitation events) are common, and δp varies systematically with mean annual temperature (Clark and
 662 Fritz, 1997; <http://isohis.iaea.org/GNIP>), δp thus broadly correlates with latitude and altitude of a site
 663 (Bowen and Wilkinson, 2002 and references therein). The global relationship between changes in δp
 664 with temperature is referred to as the ‘Dansgaard relationship’ (after Dansgaard, 1964).
 665



666
 667

668 **Fig. 6.** Overview of factors that can influence the isotope composition of a lacustrine carbonate or biogenic silica
 669 ($\delta^{18}\text{O}$). For equilibrium mineral precipitation, the oxygen isotope composition of the mineral is controlled only by
 670 the temperature and by the isotope composition of the lake water from which the mineral precipitated. If a mineral
 671 is precipitated in isotope equilibrium, paleotemperature equations can be used to estimate past temperatures and
 672 their changes. Other factors, such as pH, mineral speciation, and rates of mineral precipitation, may affect the
 673 fractionation relationship. (From Leng and Marshall, 2004).

674

675 In many lacustrine environments it cannot be assumed that either the modern or past $\delta^{18}\text{O}_{\text{lakewater}}$,
 676 reflect that of mean annual precipitation. The residence time of water in the lake and modification of
 677 water compositions by catchment and lake processes are particularly important to consider as
 678 evaporation will affect the water composition. The size of a lake in comparison to its catchment is
 679 important because the isotope composition of rain and snowfall are very variable on short time scales: A
 680 lake therefore needs to be big enough and well enough mixed for its isotope composition to ‘average out’
 681 the short-term variation and reflect mean annual precipitation. The greatest degree of variation in the
 682 isotope composition of rainfall occurs on the time scale of hours to days. Seasonal variation in

683 precipitation is likely to be much more significant in small, short residence time lakes as these tend to
684 have $\delta^{18}\text{O}_{\text{water}}$ values that are regularly displaced by later precipitation. However, if the inference of intra-
685 seasonal fluctuations is of interest, growth increments analyses of the isotope composition of molluscan
686 shells will often provide intraseasonal information (Kaandorp et al., 2005; Taft et al., 2012; Leng and
687 Lewis, 2014).

688 At temperate latitudes, monthly mean rainfall $\delta^{18}\text{O}$ typically have an overall range of 2–8‰. The
689 range increases with continentality of the site. At such sites, winter rainfall has significantly more
690 negative $\delta^{18}\text{O}$ than its summer equivalent. If a lake is very small in relation to its catchment (with
691 residence times of < 1 year), winter rainfall will be physically displaced by summer rainfall and thus
692 $\delta^{18}\text{O}_{\text{water}}$ will be influenced by seasonal variation. Therefore, deep ancient lakes with their long water
693 residence time are particularly useful for isotope studies. The precise isotope composition of lake water
694 at any time will depend on the amount of rain in the different seasons and the degree of mixing of winter
695 and summer rainfall. Even in lakes with relatively long overall residence times, surface waters may have
696 isotope compositions that reflect summer rainfall rather than mean annual precipitation if the waters
697 become stratified.

698 The $\delta^{18}\text{O}$ of components within lake sediments contain information on temperature, and there are
699 many empirically derived 'paleotemperature equations'. During equilibrium precipitation, carbonates for
700 example, have $\delta^{18}\text{O}$ compositions that decrease by about 0.24‰ for each 1°C increase in temperature
701 (Craig, 1965). There is a number of paleotemperature equations for the equilibrium precipitation. For
702 example, Kim and O'Neil (1997) proposed an equilibrium fractionation relationship, which was
703 reordered by Leng and Marshall (2004). Assuming that equilibrium precipitation has occurred, the
704 interpretation of oxygen isotope data in terms of paleotemperatures requires an understanding of two
705 processes that have opposing effects on the composition of a carbonate or silicate precipitate. At
706 intermediate and high latitudes, the oxygen isotope composition of mean annual precipitation correlates
707 directly with change in temperature with a gradient of approximately +0.6‰ °C⁻¹ (Dansgaard, 1964).
708 Opposing this effect from being transferred directly into the sediment record, the equilibrium isotope
709 fractionation between carbonate and water has a gradient of around -0.24‰ °C⁻¹. For many lake records,

710 the carbonate response to temperature will be dominated by the change in the isotope composition of
711 precipitation and effectively ‘damped’ by the opposing effect of mineral-water fractionation. In this case,
712 the measured carbonate values will covary with temperature – with an increase of $\sim 0.36\text{‰ }^{\circ}\text{C}^{-1}$ (Eicher
713 and Siegenthaler, 1976). This is reasonable for the paleoclimatic interpretation of many lakes but it
714 implicitly assumes that $\delta p/dT$ always changes according to the Dansgaard relationship.

715 Evaporation has a major influence on the oxygen isotope composition of any standing water body.
716 For closed (terminal) lakes, where water loss is mainly through evaporation, lake waters tend to have
717 high $\delta^{18}\text{O}$ (Fig. 6). Measured $\delta^{18}\text{O}$ (and δD) values are always higher than those of average precipitation
718 as the lighter isotopes of ^{16}O and (^1H) are preferentially lost to evaporation. Isotope records from such
719 lakes show large swings in composition as the ratio of the amount of precipitation to evaporation (P/E)
720 change with climate. Any interpretation of the isotope records from a lake must take into account the
721 hydrology of the lake and likely changes in hydrology that may have occurred in the past.

722 Endogenic carbonates are still the most commonly utilized materials for stable isotope analysis.
723 Endogenic carbonates are formed by photosynthetic utilization of CO_2 and resultant calcium carbonate
724 supersaturation (often mistakenly referred to as authigenic carbonate). In most temperate and high-
725 latitude regions, endogenic carbonates are precipitated mainly in the summer months during periods of
726 maximum phytoplankton productivity (Leng et al., 1999b, Teranes and McKenzie, 2001). In mid-latitude
727 and tropical lakes, phytoplankton growth may occur throughout the year although other mechanisms,
728 such as supersaturation, may also cause continuous carbonate precipitation (e.g., the Dead Sea aragonite;
729 Niemi et al., 1997). More commonly in the tropics, carbonate precipitation is related to phytoplankton
730 blooms associated with annual lake-water mixing and nutrient availability (*sensu* Lamb et al., 2002). The
731 advantage of using endogenic carbonate in stable isotope studies is that it provides an integrated climate
732 signal for the whole sample, which may be time-averaged over several years, depending on
733 sedimentation rate. However, there are some potential problems. It can be difficult to distinguish
734 between authigenic (diagenetic carbonate formed within the sediment) and allogenic (detritus derived
735 from the terrestrial environment) carbonates, especially in karstic regions, so there is always the
736 possibility of contamination of the isotope signal from a washed-in component (cf. Hammarlund and
737 Buchardt, 1996). Some carbonates do not precipitate in equilibrium with their environments.

738 Disequilibrium effects (often called ‘vital effects’ in biogenic materials) have been attributed to rates of
739 precipitation, pH effects, incorporation of metabolic fluids, and growth in microenvironments not typical
740 of the water body as a whole (Leng and Marshall, 2004). Also, there are several endogenic carbonate
741 minerals that could precipitate out in a lake and each mineral has its own mineral-water fractionation
742 (see above). In freshwater systems, calcite (CaCO_3) usually forms. However, with increasing
743 evaporation, other forms can occur, such as aragonite (CaCO_3) and dolomite ($\text{CaMg}(\text{CO}_3)_2$) (e.g., Lake
744 Bosumtwi; Talbot and Kelts, 1986). Physically separating different carbonate minerals is not easy,
745 although respective protocols are available (e.g., Dean et al., 2015). Diatom silica $\delta^{18}\text{O}$ is being
746 increasingly utilized in paleolimnology, and many of the issues are the same as with carbonates.
747 However, the main consideration is that almost pure diatom samples are required as extraction
748 techniques will liberate oxygen from all the components in the sediment (Leng and Barker, 2006).

749 Overall, isotope geochemistry is an essential part of paleoclimatological and (paleo-) limnological
750 research. Over the last few decades, the analysis of isotopes in carbonate materials has contributed
751 significantly to our understanding of a broad array of environmental change research. Nevertheless,
752 advances are still to be made. These include better preparation, analysis, and interpretation of some of
753 the less routine materials (other than carbonate), such as biogenic silica and chironomid chitin (Leng and
754 Henderson, 2013). There is also the up and coming field of determining the excess of ^{13}C – ^{18}O bonds in
755 clumped isotopes (Leng and Henderson, 2013). Bulk carbon and nitrogen isotope ratios within organic
756 materials are also frequently analyzed but do not provide specific information and, therefore, are not
757 further discussed here. Moreover, paleoenvironmental studies targeting paleohydrology and biome
758 dynamics increasingly use carbon and hydrogen isotope ratios of specific organic sources and
759 compounds, such as leaf waxes and algal lipids (for details see chapter 2.2.4).

760 2.2. *Biological and paleontological data and methods*

761 2.2.1. *Macrofossils*

762 Macrofossils include all remains of organisms from the remote past large enough to be visible
763 without a microscope. Although fossils are widely reported from lake sediment cores, obtaining

764 macrofossils is often a hope, though rarely an objective of drilling campaigns. The larger individual
765 fossils are, the smaller the chance that such fossils will be picked up in a sediment core. The diameter of
766 drill cores (typically 48–85 mm) sets a strong constraint on the possible recovery of macrofossils and
767 their size. Even fossils with a maximum linear size of 2–4 cm unlikely fall within the categories of
768 exquisite preservation and complete recovery. An additional difficulty for the recovery of macrofossils is
769 that cores are typically cut and observed perpendicular to bedding planes, so that macrofossils may easily
770 be overlooked. Recovery of macrofossils from a drilling core may also require disturbing the sediments
771 more than is desirable at the early stages of core documentation (see section 2.1.2.). To overcome this
772 issue, CT- or radiographic images can be obtained of core sections during the core documentation
773 procedure just after opening a core.

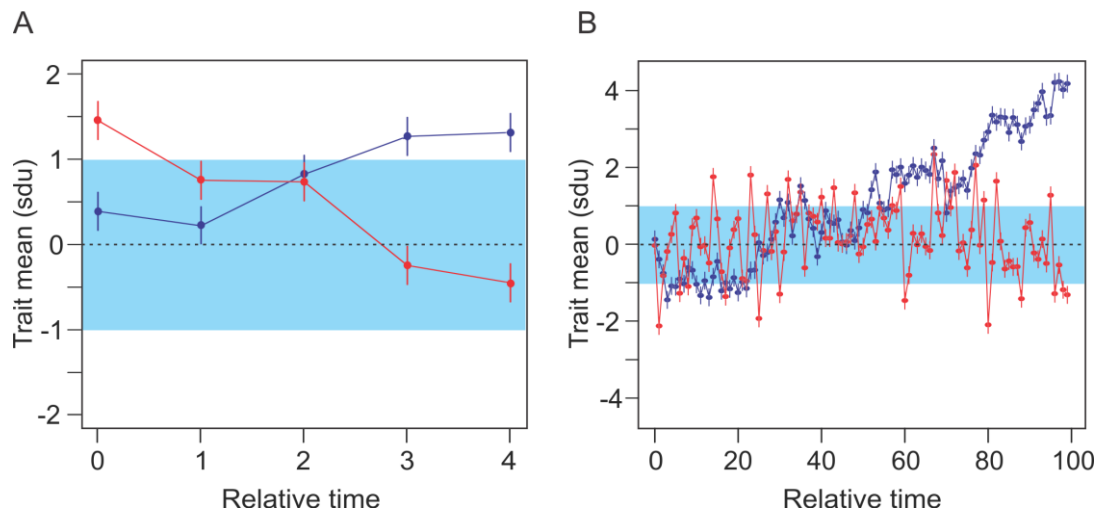
774 Macrofossils may be obtained from sediment cores as complete body fossils, such as mollusks, plant
775 leaves, or isolated biomineralized parts, such as bones, teeth, charophyte oogones, seeds, plant cuticles,
776 and wood fragments (Taviani et al., 2000; Reinthal et al., 2011; Cohen, 2012; Pepe et al., 2013, Mazzini
777 et al., 2015). Lake sediments regularly yield fossils with exceptional preservation (e.g., Richter and
778 Wedmann, 2005; Smith, 2012), and sometimes the temporal resolution is extraordinary, i.e., centennial
779 to even annual scales (Bell et al., 2006; Hunt et al., 2008; Van Bocxlaer and Hunt, 2013). The choice of
780 drilling sites affects the chances to recover macrofossils as well as fossil preservation (reviewed in
781 Cohen, 2012). For example, drilling in the deeper waters of a meromictic lake is unlikely to yield many
782 macrofossils because the benthic habitats at these sites can usually be expected to have been inhabitable
783 for benthic life for most of the lake's history (see section 2.1.1.), although remains of pelagic organisms,
784 such as lake sardines (Cyprinidae) may end up in such cores (Reinthal et al., 2011). Hence, macrofossils
785 from nearshore habitats generally can be found only in low abundances in offshore cores. Drill sites
786 relatively close to the shores, in general, provide greater potential to recover macrofossils (Harzhauser et
787 al., 2013). However, such sites usually contain large proportions of sand and gravel (indicative for
788 greater water energy and poorer preservation potential), and are subject to sediment redeposition, which
789 makes them undesirable targets for paleoclimate studies and difficult to drill (see also section 2.1.1.).
790 Mid-lake topographic highs are often selected as disturbance can be anticipated to have been limited, but
791 such sites are unlikely the most suitable target of drilling from the perspective of macrofossil recovery. If

792 recovered, macrofossils may serve several research goals, and because of the practical limitations to find
793 macrofossils, we discern here between goals that can be reached with a low number of specimens and
794 those that require the sampling of larger numbers of macrofossils.

795 Single or scarce macrofossil finds may provide valuable information for the study of
796 paleoenvironments or organismal evolution. They can also facilitate dating analyses, either directly if the
797 fossil provides datable substances (e.g., charcoal, shells, seeds of terrestrial plants, etc.), via the study of
798 ancient amino acids (see section 2.2.5.), or if the macrofossil can be placed into a biostratigraphic
799 framework for the studied basin. It can also be used indirectly if the fossil can be inserted with
800 reasonable constraint into a time-calibrated phylogeny of extant taxa (see section 2.2.9.). Moreover,
801 macrofossils may provide unique opportunities for paleolimnological reconstructions, either via isotope
802 studies (see section 2.2.8.) or via reconstructions directly based on the fossils' properties or habitat
803 characteristics (see section 3.1.). Beyond calibrating the molecular clock (see section 3.2.), macrofossils
804 may give a good insight into organismal evolution including information on how long the taxon is
805 present in the basin, on ancestral character states (see section 3.4.), and on morphological adaptation by
806 comparing the morphology of the fossil to that of modern populations. In the most fortunate case, the
807 fossil may be exceptionally preserved so that attempts to isolate ancient amino acids or DNA can be
808 undertaken (see sections 2.2.5., 2.2.6.).

809 However, most biological goals relate to community changes over time or morphological evolution,
810 and thus require sampling multiple taxa or individuals per interval (also see section 3.1.). As indicated
811 above, chances to recover macrofossils in substantial numbers relate first to the size of such fossils, but
812 furthermore also to preservation potential (e.g., CaCO₃-bearing fossils will corrode when deposited
813 below the local carbon compensation depth), general abundance, and lifestyle (e.g., benthic vs. pelagic).
814 Therefore, most of the macrofossils that can be used for goals requiring abundant finds are only
815 marginally larger than microfossils (e.g., fish teeth, or small mollusks; see section 2.2.2.). Continuity of
816 the recovery of such fossils throughout a core relates primarily to the preservation potential under
817 changing chemical conditions (dissolution, corrosion, and abrasion diminish preservation potential) and
818 the patchiness of a taxon's occurrence. Patchy taxa can be expected to occur with great numbers at few
819 intervals, whereas more evenly distributed taxa would occur in smaller numbers throughout more

820 intervals within a sediment core. The abundance of the target taxon per interval of sediment is the main
 821 determinant of the resolution at which the taxon's macrofossils can be studied. For example, if an
 822 interval of 2 cm of sediment yields on average 40 seeds, but only 2 fish teeth and 4 mollusks, then the
 823 core can be analyzed in 2 cm intervals for plant remains, but perhaps only in 10+ cm intervals for the
 824 study of fish teeth and mollusks, which affects the power to discover trends (see the simulation in Fig. 7).
 825



826
 827
 828 **Fig. 7.** The feasibility to detect ecological or evolutionary signals from drilling data depends on the length of a time
 829 series (here expressed as relative time). Simulated time series of a morphological trait (scaled in within-sample
 830 standard deviation units; sdu) are represented in A) and B). The trait indicated in blue was simulated with a model
 831 of directional change, the red trait with a model of morphological stasis. Parameters were identical for each model
 832 in A and B. Only over an extended period of time (B), morphological stasis and directional change can be
 833 distinguished (blue area = 2 sdu). Modeling was performed in R 3.3.1 (R Development Core Team, 2015) with the
 834 paleoTS package (Hunt, 2006, 2012).
 835

836 Upon detecting macrofossil remains in lacustrine sediment cores, the specimens are usually picked
 837 out, identified and/or counted prior to further analyses, such as dating or the study of biomarkers, isotope
 838 and element geochemistry. The advantage of using macrofossils for the latter studies is that analyses may
 839 be performed on a single fossil and may yield data on, e.g., seasonal fluctuations. This procedure allows
 840 obtaining a more accurate signal than when individuals are pooled or a sample of bulk organic remains is
 841 analyzed, because these latter samples provide averaged signals.

842 Identifications of macrofossils may be challenging because taphonomic processes, such as time

843 averaging, sorting, and post-mortem pooling (e.g., deposition of two taxa with fine-scaled habitat
844 differences in the same assemblage), might considerably alter the amount of variation observed in a
845 fossil assemblage in comparison to that in a modern population (e.g., Bell et al., 1987; Bush et al., 2002).
846 Furthermore, chronospecies or taxa with character states intermediate to those of two or more modern
847 species remain a problem. They result in the poor applicability of identification keys and potentially in
848 doubtful identifications—much of these aspects are inherent to fossils and hence, apply for microfossils
849 as well. Initial exploratory analyses may be conducted to examine occurrence data (counts) belonging to
850 different groups of taxa (e.g., endemics vs. non-endemics), or to calculate biodiversity or community
851 estimators/indices to compare a number of target assemblages (e.g., faunal and floral compositions
852 before, during, and after a climatic or geological event). If qualitative observations suggest potential
853 morphological changes, measurements may be taken to accurately document these changes through time.
854 Such measurements can range from traditional caliper-measurements to studies of size and shape with
855 fractal dimensions or geometric morphometrics. The choice for a particular method typically depends on
856 the complexity of the signal, the time required to document/measure a single specimen, and the total
857 number of specimens to be studied. ‘Targeted macrofossils’ may also be studied with more time-
858 consuming 3D scanning methods and/or biogeochemical analyses (e.g., to document chemical
859 composition, to study diagenetic processes, or to get information on an environmental proxy).

860 After exploratory analyses, more in-depth statistical and time-series studies can be undertaken.
861 Fossils (mainly microfossils) encountered in sediment cores are regularly used for analyses of
862 community composition, often in relation to environmental change (e.g., Cohen et al., 2007; Kröpelin et
863 al., 2008; Harzhauser et al., 2013; also see sections 2.2.2., 2.2.3.) or to document morphological change
864 and evolution in the fossil record (e.g., Pearson and Ezard, 2014 and references therein). For studying
865 shape/community changes related to environmental change or organismal evolution, statistical and time-
866 series analyses provide a useful framework to explore and test relationships between predictor variables
867 and organismal change, or to fit models of morphological evolution. Major determinants of the power of
868 such time-series approaches to discover ecological patterns will be the strength of the association
869 between the predictor variables and organismal/community change, the range of values of the predictor
870 variable observed throughout the core, and the variation in the dependent variable. To reliably document

871 patterns of morphological evolution in a fossil lineage, a major determinant of analytical power will be
872 the ratio of variation within individual samples to the changes between consecutive samples. For
873 example, fewer specimens per interval will be required to document a strong morphological trend in a
874 time series with limited within-sample variation than for a time-series with the same trend and great
875 within-sample variation. Additionally, as mentioned, the number of sampled intervals throughout the
876 core for which all required data are available and, hence, the length of the time-series is a determining
877 factor of analytical power (Fig. 7). To assess the feasibility of time-series analyses with drill core data,
878 detailed analyses of the anticipated patterns of change and the various components that contribute to the
879 variation in fossil assemblages may be required. Explorations examining variation in homologous or
880 analogous modern faunas and floras may be required to understand how different sources of variation
881 contribute to the total variation observed in fossil assemblages. Dieleman et al. (2015) presented such an
882 exploration for the study of fossil cichlid teeth from the African crater Lake Challa.

883 Offering suggestions on how future lake drilling campaigns could be designed optimally for the study
884 of macrofossils is not easy. First, strategies depend on the group of macrofossils that is specifically
885 targeted. Second, each ancient lake has unique ecosystems that differ from those of other lakes and the
886 design of the program needs to be adjusted to the specific target lake. Third, the multidisciplinary nature
887 of many drilling operations may regularly weaken the feasibility of obtaining a continuous fossil record
888 due to compromises in site selection (Cohen, 2012; also see section 2.1.).

889 Nevertheless, promising subprojects based on macrofossils can be constructed, and the following
890 aspects can augment the potential of such programs. Drilling campaigns usually aim to retrieve multiple
891 cores and one (or some) of the drilling locations may be selected where abundant macrofossils can be
892 expected (e.g., Wagner et al., 2014a). Moreover, fossil-bearing outcrops may be present in the basin, and
893 these too may contain a rich archive of faunal and environmental history. An outstanding example is the
894 mollusk fauna from the Turkana Basin (Williamson, 1981, 1985; Van Bocxlaer et al., 2008). However,
895 this example also highlights that interpretations of outcrop data are highly dependent on the available
896 stratigraphic control. Integrating data from lake drilling and the study of fossiliferous outcrops would be
897 a powerful approach to eliminate the weaknesses of individual strategies, and hence to solidify the
898 documentation of the lake's natural history. Specific hypotheses that were constructed based on the study

899 of the modern fauna and that are testable with macrofossils from sediment cores can provide good
900 objectives for a drilling strategy that maximizes the potential to recover macrofossils.

901 Several future trends in data generation and analyses specifically geared to macrofossils are
902 conceivable. An important starting point is the diameter of the core. Bigger diameters would increase the
903 potential to obtain macrofossils, however, they would also increase the chance of technical difficulties.
904 Core scanning (i.e., CT-scanning) in a way that allows the non-invasive detection of various
905 macrofossils on bedding planes based on biomineral concentrations, density contrasts, or shape would be
906 invaluable to select targeted parts of the cores for detailed examination, with minimal disturbance to
907 other parts. Furthermore, more automated methods in fossil retrieval and visualization would help and
908 speed up the development of goals based on large numbers of fossils. Automated identification,
909 measurement, and digitization would greatly facilitate the study of macrofossils, but the desirability of
910 such developments ranges beyond the drilling communities at large (e.g., Houle et al., 2003).

911 2.2.2. *Microfossils*

912 Lake sediments are also valuable archives of microfossils, i.e., remains or traces of organisms from
913 the past that require a microscope for study. Typical microfossils in sediment cores include plant pollen
914 from seed plants, single-celled protists with biomineralized shells (e.g., calcareous dinoflagellate cysts
915 and siliceous diatoms), as well as small shelled invertebrates, such as ostracods and micromollusks. As
916 the latter share many characteristics with macrofossils (see section 2.2.1.), this review focuses mainly on
917 the discussion of pollen and diatom microfossils, which have great potential for environmental studies
918 based on lake-cores.

919 920 *Diatoms*

921
922 Diatoms (phylum Bacillariophyta) are photosynthetic eukaryotes. Their silica shell ('frustule')
923 preserves exceptionally well, making diatoms one of the most abundant microfossil taxa in lacustrine
924 systems (Gross, 2012). Moreover, the frustules often display species-specific ornamentation, providing
925 important information for identification and classifications (Round et al., 1990).

926 Ancient lakes regularly display an exceptional diatom biodiversity (Cocquyt, 1998; Levkov et al.,

927 2007; Mackay et al., 2010; Kulikovskiy et al., 2012; Wagner et al., 2014a) with the total number of
928 morphotypes often exceeding 1,000 (Levko and Williams, 2012). Most diatom species are benthic; only
929 few species are euplanktonic and spend their entire life cycle in the water column (Stoermer and Edlund,
930 1999; Khursevich and Prokopenko, 2009; Snyder et al., 2013; Recasens et al., 2015).

931 Current diatom research in ancient lakes strongly focuses on the reconstruction of paleoenvironments
932 and past environmental fluctuations. One of the most frequently used approaches is the chronological
933 assessment of changes in species composition and relative abundances of the dominant species at various
934 temporal scales (orbital, millennial, centennial, or pluriannual). These community studies are typically
935 undertaken at low taxonomic resolutions (i.e., species from one or several genera are combined) and with
936 a relatively large number of valves to be counted per sampled sediment interval (e.g., 400–600). As
937 habitat characteristics and ecological preferences of many diatom species are well characterized, the
938 collective diatom community found in a sediment sample then becomes an important source for analysis,
939 e.g., with paleoecological transfer functions (see section 3.1.).

940 However, given that diatom communities in ancient lakes are often dominated by endemic species,
941 approaches based on transfer functions may be problematic due to the lack of analogues in the existing
942 calibration sets. In fact, given the old age of many diatom fossils in ancient lakes, they may not even
943 have extant analogues in the lake (= supralimital evolution; Wesselingh, 2007; Mackay et al., 2010).
944 Nonetheless, some attempts have been made to establish modern diatom calibration sets for ancient lakes
945 (e.g., Lake Baikal; Mackay et al., 2003). These attempts try to overcome the limitation of missing
946 modern analogues by examining modern species compositions along environmental/spatial gradients
947 within the lake. Moreover, some studies suggest that species with similar size and comparable
948 ultrastructural features have similar ecological preferences (e.g., Winder et al., 2009), and this
949 assumption may also be used in establishing calibration sets. In combination with other proxies (e.g.,
950 biogeochemical data derived from the same sediment record), these new approaches enable a more
951 reliable paleoenvironmental interpretation of diatom communities (Bradbury, 1999).

952 Another focus of research on diatoms obtained from ancient lake sediment cores is the assessment of
953 a lake's primary productivity (*sensu* Battarbee, 1986). As diatoms are important primary producers, their
954 concentration is a good indicator for productivity and epilimnetic nutrient availability (Zhang et al.,

955 2016). Productivity data have been used, for example, to identify successive glacial and interglacial
956 stages (Rioual and Mackay, 2005; Melles et al., 2012; Meyer-Jacob et al., 2014b). This approach is
957 especially promising if environmental fluctuations are not anticipated to have caused significant changes
958 in diatom species composition (e.g., Cvetkoska et al., 2015).

959 However, diatom concentrations in drill cores may be affected by taphonomic processes, such as
960 differential dissolution (Ryves et al., 2006). Parameters, such as low temperature, high pH, high detrital
961 mineral content, and grazing, may (selectively) decrease valve preservation, and can lead to the loss of
962 specific taxa, morphological details on individual frustules, or even the entire assemblage (Mackay,
963 2007; Reed et al., 2010). Moreover, biases in diatom concentrations may result from inwash of
964 specimens from rivers, or from the loss of specimens via the lake's outflow (O'Sullivan, 2004). In
965 addition, diatom concentrations do not account for size differences in taxa or changes in sediment
966 accumulation rates. Therefore, the biovolume accumulation rate, i.e., diatom concentration divided by
967 sediment accumulation rate and corrected for size differences between taxa (O'Sullivan, 2004), could be
968 a more powerful proxy for a lake's productivity (Rioual and Mackay, 2005).

969 Several other aspects complicate interpretations of diatom records from ancient lakes (Bradbury,
970 1999). Besides the taxonomical and ecological uncertainties mentioned above, only a fraction of the
971 actual (endemic) biodiversity is known. Moreover, recent changes in diatom species and genus concepts
972 have led to an introduction of more than 5,000 new names (Fourtanier and Kociolek, 2011). Even
973 specialists with several years of experience have difficulties keeping up with these changes and, more
974 importantly, to unify their taxonomic opinions with those of colleagues. Finally, diatom analyses are
975 time consuming, thus limiting the number of samples that can be analyzed. Therefore, samples from
976 sediment cores are often subdivided and analyzed in parallel by several researchers, creating a potential
977 bias in observations. Moreover, species within common planktonic genera, such as *Cyclotella* and
978 *Stephanodiscus*, differ in morphological features that are difficult to assess with a light microscope,
979 potentially increasing determination errors. The problem of taxonomic accuracy is further increased by
980 the cryptic nature of some closely related diatom species, on the one hand, and by a partially high
981 phenotypic plasticity, potentially driven by environmental factors, on the other hand (e.g., Cvetkoska et
982 al., 2012; García et al., 2012). These problems are progressively addressed via molecular analyses and by

983 cultivation experiments, which have shown that diatoms can change morphological features during
984 cultivation (e.g., Abarca et al., 2014; Rose and Cox, 2014). Nonetheless, diatom studies can provide very
985 valuable insights into the primary productivity of ancient lakes, and into the interpretation of past
986 environmental changes (see section 3.1.).

987

988 *Pollen*

989

990 Plant pollen produced by terrestrial seed plants is frequently deposited in lacustrine systems by wind
991 ('pollen rain'), riverine inflow, and animals (e.g., O'Sullivan, 2004). Similar to diatoms, pollen is often
992 well-preserved in long sediment cores. The high resistance to degradation under anoxic/hypoxic
993 conditions is due to the polymer sporopollenin, which is a component of the outer wall (= exine) of
994 pollen grains (Brooks and Shaw, 1978).

995 The pollen deposited in sediment cores from ancient lakes is an important archive of past
996 environmental and climatic changes. Pollen is typically present throughout the sediment core, thus
997 enabling continuous analyses over long time scales. Moreover, their deposition is usually only affected
998 to limited extent by hydrological and chemical changes within the lake, and, hence, pollen records
999 typically provide accurate reflections of terrestrial events near the studied lake (O'Sullivan, 2004).

1000 Finally, pollen deposition in sediment cores may permit the reconstruction of regional changes as pollen
1001 rain is often dispersed over long distances. Thus, that the record may contain the signal of both local and
1002 regional drivers of environmental change (e.g., Wagner et al., 2014a).

1003 As the composition and provenance of pollen in lake cores are often complex due to varying pollen
1004 productivity and dispersal rates (Faegri et al., 1989), significant efforts have been made to understand the
1005 extent to which these records represent past vegetation covers (e.g., Davis et al., 2013; Schüler et al.,
1006 2014; Trondman et al., 2015; see also section 3.1.) and, hence, how such records should be interpreted.

1007 Pollen is generally identified to the genus level, because the pollen of many closely related species
1008 cannot be distinguished using transmission light microscopy. Subsequent paleoreconstructions are
1009 generally done using pollen diagrams, which utilize information on pollen composition, concentration,
1010 and influx values (Berglund and Ralska-Jasiewiczowa, 1986). Past vegetation covers and floral
1011 compositions can then be related to the underlying climatic drivers. Tzedakis et al. (1997), for example,

1012 observed a close correlation between herbaceous vegetation and ice volume at a global scale during
1013 glacial intervals, whereas forest physiognomy and development appeared closely related to changes in
1014 temperature and humidity during interglacials. This study and following work (Tzedakis et al., 2001)
1015 reinforced evidence for a broad correspondence between climate signals provided by pollen data in long
1016 lacustrine records and oxygen isotopes from marine cores. Other pollen records have allowed to link
1017 short-term vegetation oscillations and centennial-scale climatic events on various smaller scales (e.g.,
1018 Dansgaard et al., 1993; Broecker, 1994; Bond and Lotti, 1995). Pollen records from cores of ancient
1019 lakes that have a well-established chronology can thus provide a better understanding of climate forcing
1020 from local to global scales over several glacial/interglacial cycles (Litt et al., 2014; Sadori et al., 2016).

1021 Other objectives of pollen analyses in long and continuous sediment records are the reconstruction of
1022 species dynamics, dating of extinction events (see also section 3.2.), and the inference of possible refuge
1023 areas. Bertini (2010), for example, could show that extinction events following climate changes did not
1024 occur synchronously across ecosystems. However, geographically-related records, in general, may show
1025 somewhat different vegetation dynamics. Part of the problem is that many previous sediment records
1026 have not been studied with a high temporal resolution and/or lack a precise chronological control.

1027

1028 Given the challenges and limitations outlined above for diatom and pollen microfossil analyses from
1029 sediment cores, the following recommendations for future deep drilling projects are suggested:

1030

- 1031 i) Microfossil studies should be hierarchically structured. As analyses from long records are very
1032 laborious and time-consuming, the first target should be to produce low/medium resolution data
1033 (e.g., skeleton pollen diagrams with key pollen curves, Sadori et al., 2016 or stratigraphic diatom
1034 diagrams, Cvetkoska et al., 2015). This enables a preliminary chronological alignment with major
1035 environmental and climatic fluctuations. Only then, high-resolution diatom and pollen studies
1036 should be conducted.
- 1037 ii) The temporal sampling design should adhere to the main question and time-scale of the respective
1038 project. However, in case of uncertainties about the temporal resolution required, subsampling should
1039 be set up in a way that samples for higher resolution studies are available even if a first analysis is to

1040 be conducted in lower temporal resolution. This is because resampling of sediment cores that are
1041 already in long-term storage may be difficult and/or costly.

1042 iii) Sediment subsampling should be coordinated (see section 2.1.4.). In order to be able to combine
1043 microfossil data sets and link them to other biotic and abiotic data generated in multidisciplinary deep
1044 drilling projects, samples should, whenever possible, be taken from the same sample depth.

1045 iv) Protocols for sample preparation and microfossil identification to be used by all biologists involved
1046 in the project have to be implemented. A ‘taxonomy working group’ should be established that
1047 elaborate and share the taxon list with all investigators, defines diagnostic characters for problematic
1048 taxa, solves taxonomic disputes, and implements quality control procedures (e.g., Munro et al., 1990).

1049 v) Microfossil studies should involve specialists in climate reconstruction and empirical modeling at an
1050 early stage to improve data quality and the power of subsequent analyses.

1051
1052 We expect to see considerable advances in microfossil biodiversity research in the future, partly
1053 driven by ongoing deep drilling projects in ancient lakes. Comparative molecular investigations of recent
1054 taxa, particular for diatom analyses, may help clarifying systematic problems, mainly in respect to
1055 cryptic species and species with high phenotypic plasticity (Kermarrec et al., 2013). Robust species-level
1056 phylogenies could then be used to statistically identify those morphological character states or
1057 combinations thereof that have a high diagnostic power. These characters could possibly also be applied
1058 to paleo-morphotypes.

1059 As to future pollen fresearch, a better link to studies of plant macrofossils could be established (Birks
1060 and Birks, 2000; see also section 2.2.1.). Macrofossils can often be identified with more taxonomic
1061 precision than pollen, and in the context of lake drilling, they may provide supplementary information,
1062 particularly on water plant communities (Birks and Birks, 2000; Sadori et al., 2010), and taxa that
1063 produce very little or no pollen. Integrated paleobotanical analyses may allow creating a more
1064 comprehensive picture of paleoenvironmental changes in ancient lakes and their watersheds.

1065 Another promising development is the 3D reconstruction of microfossils. Whereas scanning electron
1066 microscopic 3D approaches are still costly and time consuming (e.g., Mansilla et al., 2015), light
1067 microscopic 3D solutions are meanwhile offered by several companies.

1068 Bioinformatic advances in image analyses have also led to improved automatic identifications of
1069 diatoms (e.g., Jalba et al., 2005; Mosleh et al., 2012; Kloster et al., 2014) and pollen (Guru et al., 2013;
1070 Holt and Bennett, 2014; Marcos et al., 2015), potentially reducing processing time and identification
1071 biases. Future improvements may further increase the performance of these approaches beyond their
1072 current application as preliminary sorting tools.

1073 2.2.3. *Subsurface biosphere*

1074 Over the past decades, marine and continental drilling projects have shown a growing interest in
1075 understanding the role of microbes in the complex chemical reactions occurring in the sediments and or
1076 the sediment/water interface (Konhauser, 2007). Microbial activity in the water column of modern lakes
1077 and oceans is well known in comparison to effects of microbial life in sediments, but despite controversy
1078 on specifics, scientists generally agree that the impact of the deep marine biosphere on global
1079 biochemical cycles is massive (Whitman et al., 1998; Kallmeyer et al., 2012). As a result, there has been
1080 a substantial increase in investigations dealing with microbial activity in freshly retrieved sediments
1081 (D'Hondt et al., 2002; Inagaki, 2010). This hidden microbial mass as well as the study of the interactions
1082 between microbes and sediments in the marine environment is known as 'deep biosphere research'.
1083 Analogously, the study of active microbes in lake sediments through scientific drilling have been labeled
1084 as 'subsurface biosphere research'. Many microbes in this subsurface biosphere grow extraordinarily
1085 slowly and under extreme conditions (Røy et al., 2012). Their study is critical to understand the
1086 physiological abilities and biogeochemical impact of subsurface life within the sedimentary column.

1087 Subsurface biosphere studies have only recently been implemented in continental deep drilling
1088 projects (Vuillemin et al., 2010; 2013a, b; 2014a; b; Glombitza et al., 2013; Thomas et al., 2014, 2015;
1089 also see Ariztegui et al., 2015 for a detailed description of the different sites).

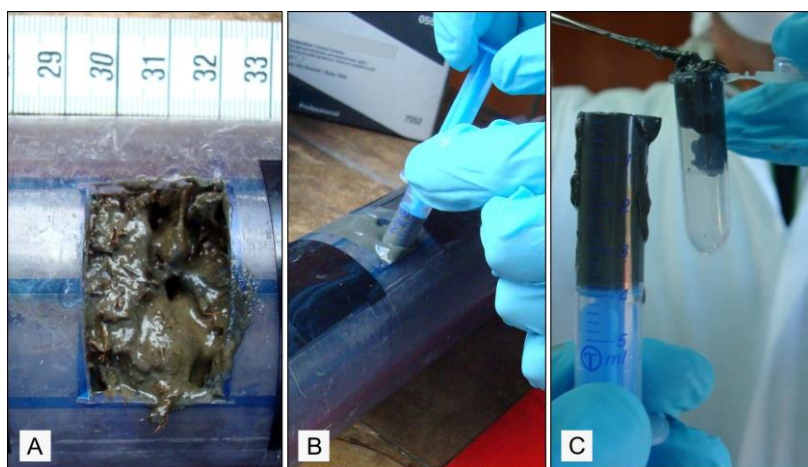
1090 The systematic study of the subsurface biosphere in lacustrine sediments will allow geologists and
1091 biologists to identify the magnitude and impact of microbes during early diagenesis on both sediments
1092 and biological remains. Geomicrobiological studies permit the identification of signatures of former
1093 microbial activity recorded in the sediments as well as investigating their impact in biogeochemical
1094 cycles. Some open questions are:

- 1095 i) What are the source, composition, and global significance of subsurface communities in lacustrine
1096 basins?
- 1097 ii) What is the impact of environmental change on subsurface biodiversity? Do changes in
1098 environmental conditions generate changes in microbiomes population diversity and density?
- 1099 iii) How does the diversity and activity of microbial life vary with depth, geochemistry, sediment
1100 composition, and age?
- 1101 iv) How does the diversity of active and non-active microbes relate spatially across the lacustrine basin
1102 at given time slices?
- 1103 v) How do microbes resolve the paucity of nutrients and energy as well as the limits of life in the
1104 subsurface of lakes under contrasting physicochemical conditions?
- 1105 vi) What is the influence of subsurface communities on paleoenvironmental and paleoclimatic proxies,
1106 minerals, and hydrocarbon reservoirs? How do they alter sediments and hence the data and
1107 interpretations of other workers on deep drilling cores?
- 1108 Cell counts in recent continental deep drilling in mesosaline Lake Potrok Aike (Patagonia, Argentina)
1109 and alkaline Lake Van (Turkey) indicate variable microbial activity at sediment depths reaching down to
1110 100 mcd (Vuillemin et al., 2010; Kallmeyer et al., 2015). Additionally, investigations in hypersaline
1111 Dead Sea sediments (Levantine region) have shown a dominant Archaea population down to 200 mcd
1112 (Thomas et al., 2014, 2015). These results indicate that further research is critical to obtain more detailed
1113 information about the fine-scale mechanisms controlling microbial life in the deep subsurface of ancient
1114 lake sediments while answering several of the aforementioned questions.

1115 Compared to their marine counterparts, the study of lacustrine sediments is often logistically easier.
1116 However, the implementation of subsurface biosphere studies in lake systems has required new sampling
1117 techniques that provide the needed aseptic conditions to avoid contamination. Whereas most ocean
1118 drilling vessels have a dedicated laboratory for geomicrobiological sampling on board, the size and
1119 configuration of lake drilling platforms prevent setting up a comparable sampling laboratory with
1120 conditions of asepsis. Therefore, special on-shore facilities are required to solve this problem. Core
1121 sections are transported to this laboratory as frequently as possible, depending on the logistics of the
1122 drilling (e.g., distance to the shore, frequency of drilling, and crew rotation; also see section 2.1.).

1123 Moreover, *in situ* sampling procedures allow recovering aseptic samples as well as determining the
1124 presence of active microbes (Fig. 8; for details see Vuillemin et al., 2010).

1125



1126

1127

1128 **Fig. 8.** Subsurface biosphere sampling from sediment cores obtained from Lake Potrok Aike (Patagonia,
1129 Argentina). A) Sampling window cut in the core under aseptic conditions. B) Methane sampling. C) DNA sampling
1130 (photo credits: A. Vuillemin and D. Ariztegui).

1131

1132 Ideally, a geomicrobiological study should be planned well ahead of the start of a drilling operation. It
1133 is preferable to have a dedicated core for microbiological and geochemical studies. This will secure the
1134 best possible sampling conditions but can cause problems of hole-to-hole correlation because the
1135 microbiological results have to be tied with other sedimentological, biological, and geochemical
1136 parameters as well as core chronologies. To overcome this issue of missing stratigraphic information,
1137 petrophysical properties (whole core MSCL; see section 2.1.4.) of cores dedicated for
1138 geomicrobiological research can be measured in the field, or borehole logging can be used (see section
1139 2.1.3). These petrophysical or logging datasets can then be used to generate at least a rough hole/core-to-
1140 hole/core correlation and help to embed the geomicrobiological datasets into the common stratigraphic
1141 framework and other continuously generated datasets from other holes/cores of the same site.

1142 In some cases it may, for logistical and/or budgetary reasons, not be possible to dedicate an entire
1143 drill core to microbiological investigations. If this is the case, there is the possibility to sample the core
1144 catchers for microbiological studies. When working with core catcher materials, special care should be
1145 taken to avoid contamination. A recent study that utilized core catcher samples for microbiological

1146 analyses from the Dead Sea ICDP sediment cores and followed appropriate sampling protocols, has
1147 proven to produce reliable results (Thomas et al., 2014, 2015).

1148 The significance and validity of the results of subsurface biosphere studies is largely dependent on the
1149 quality, speed, and prevailing conditions of the initial sampling. Hence, it is critical to attain a
1150 methodological standardization for all lake and ocean deep drilling sites in order to be able to compare
1151 results between different campaigns and environments. A generalized protocol would allow reducing the
1152 impact of contamination issues, determining the best method to accomplish on-site cell counting,
1153 choosing the appropriated sampling methods for further molecular characterization, and designing a
1154 proper strategy for sample archiving. Due to the different nature of each drilling project, a protocol for
1155 standardized biological sampling, processing, and analysis would be a significant accomplishment.
1156 Recently, the development of genomics and the emergence of high-throughput DNA sequencing
1157 technologies have been opening up new possibilities including the expansion of databases, which contain
1158 crucial information to define the metabolic pathways of different microbes (also see section 2.2.5.). The
1159 latter combined with laboratory culture experiments will be critical to constrain the impact of active
1160 microbes on, for example, the carbon cycle and diagenetic processes in the sediments.

1161 As in ocean research, a main challenge for the development of subsurface biosphere studies in
1162 lacustrine settings is to communicate the potential of these investigations to the broader scientific
1163 community participating in deep drilling projects. It is also important to involve geomicrobiologists in
1164 discussions about the effects of an active biosphere on the subsurface environment and sediment
1165 composition. Paleoclimatological reconstructions largely depend on the use of a variety of petrophysical,
1166 geochemical, and biological proxies. Proxy responses are typically interpreted to reflect the
1167 environmental/climatic conditions during the time of deposition. However, organic compounds (ancient
1168 DNA, see section 2.2.5; biomarkers, see section 3.3.) as well as element and isotope compositions (see
1169 section 2.2.8.) may be altered by microbially induced processes long after deposition, thus biasing and
1170 complicating the interpretation of proxies. A precise understanding of the influence of different
1171 microbial communities on compositional changes after burial is of vital importance for a better
1172 understanding of proxy responses and their interpretation.

1173 2.2.4. *Biomarkers*

1174 Biomarkers are source-specific organic molecules, i.e., they are synthesized by living organisms in
1175 aquatic or terrestrial settings for specific organic tissue types, e.g., cell membranes or protective and
1176 supporting tissue, or to fulfill specific functions, e.g., energy storage or pigmentation. The presence of
1177 such compounds in lacustrine sedimentary records is a testimony that, in the past, certain groups of
1178 organisms occurred in aquatic and terrestrial habitats of the catchment of the studied lake basin. In
1179 ancient lakes, occurrences of organisms may even be modified by geological processes, such as tectonic
1180 subsidence or uplift, changing catchment topography, and material fluxes within the watershed. This, in
1181 turn, can be detected by biomarkers. Very few individual biomarkers are species-specific, and there is
1182 considerable overlap in biomarker profiles of large groups of organisms. Still, many biomarkers can be
1183 chemotaxonomically assigned to groups of organisms that characterize a specific habitat. Furthermore,
1184 organic matter pools of living and decaying biomass (aquatic/terrestrial vegetation, sinking
1185 particles/plant litter, soil organic matter) produce equally specific combinations of biomarkers, i.e., an
1186 organic geochemical fingerprint that can be interpreted to represent an ecosystem (e.g., Holtvoeth et al.,
1187 2016). Thus, ‘source-specific’ may refer to a specific organic matter pool, a group of organisms, or to
1188 individual species. The most fundamental distinction as to the sources of organic matter in sedimentary
1189 records that biomarkers can provide is between aquatic and terrestrial plant biomass, based on the fact
1190 that vascular plants (higher land plants) require structurally supportive and protective tissues that are not
1191 present in non-vascular, aquatic plants. Some other biomarkers indicate highly specific adaptations of their
1192 source organisms to environmental conditions. For example, pigments of anaerobic phototrophic bacteria
1193 indicate past anoxia in the photic zone of the water column when found in lacustrine sediments (Hanisch
1194 et al., 2003; Castañeda and Schouten, 2011).

1195 The association of biomarkers to specific ecological functions or conditions highlights a fundamental
1196 principle behind many biomarker applications in paleo-environmental research. Organic matter
1197 inventories over time may document fluxes in various biomarkers in response to largely climatically
1198 controlled environmental parameters, e.g., temperature and the supply of moisture. Factors, such as
1199 catchment topography and lake bathymetry, also determine organic matter pools through, for example,

1200 soil thickness and stability, weathering and erosion rates, run-off modes, the extent of the littoral, all of
 1201 which also affect organic matter degradation during storage, transport, and deposition. Thus, the two
 1202 main approaches in biomarker studies are to infer i) sources of organic matter and ii) environmental
 1203 parameters (Table 2.)

1204

1205 **Table 2**

1206 Biomarker-based approaches that determine sources of organic matter and environmental parameters, with
 1207 examples of relevant literature.

1208

Parameters	Molecular evidence	References
Organic matter source		
Aquatic biomass (phytoplankton/algae, zooplankton, macrophytes, bacteria, archaea)	- chemotaxonomic compounds, incl. membrane lipids, storage lipids, pigments - compound-specific stable isotopes ($\delta^{13}\text{C}$, $\delta^2\text{H}$, $\delta^{15}\text{N}$)	Volkman et al. (1998), Meyers (2003), Castañeda and Schouten (2011)
Terrestrial vegetation (woody/non-woody vegetation, angiosperms/ gymnosperms, C3/C4 plants)	- chemotaxonomic compounds and compound distributions, incl. membrane lipids, leaf waxes, lignin phenols - compound-specific stable isotopes ($\delta^{13}\text{C}$, $\delta^2\text{H}$, $\delta^{15}\text{N}$)	Meyers (2003), Castañeda and Schouten (2011)
Soil organic matter (belowground biomass, incl. degraded plant debris, roots, bacteria, archaea, fungi)	- biomarkers from soil microbial organisms (membrane lipids) - biomarkers from root material (suberin monomers)	
Burned biomass	- pyrogenic compounds	Denis et al., 2012
Environmental parameters		
Lake surface temperature (LST)	- alkenone unsaturation index (Uk_{37} , from C_{37} alkenones of haptophyte algae) - tetraether index (TEX_{86} , from glycerol dialkyl glycerol tetraethers/GDGTs of aquatic archaea)	Castañeda and Schouten (2011); Schouten et al. (2013)
Mean annual air temperature (MAT)	- methylation and cyclisation of branched archaeal tetraethers (MBT, CBT)	
Moisture source (hydrology)	- compound-specific hydrogen isotopes ($\delta^2\text{H}$)	Huang et al. (2002, 2004), Sachse et al. (2004)
Soil pH	- MBT, CBT	Weijers et al. (2007), De Jonge et al. (2014)

1209

1210 Both approaches are frequently applied in paleoenvironmental studies as ecosystem functions are
1211 adjusted to environmental parameters. In tropical and subtropical settings, for example, information on
1212 moisture supply and evapotranspiration, which can be detected by plant wax $\delta^2\text{H}$, can be combined with
1213 carbon isotope data of the same plant wax $\delta^{13}\text{C}$ (e.g., Berke et al., 2012) to trace the abundance of C4
1214 vegetation through time. Carbon and hydrogen isotope records indicating shifts in climatically controlled
1215 hydrology and types of organic matter can then be used for climate modeling (Aichner et al., 2015).

1216 Lipid biomarkers provide highly specific proxy data that support and validate data from other organic
1217 sediment components, in particular, palynological and bulk organic matter data (see section 2.2.2.).
1218 Relatively fast and cost-efficient bulk geochemical proxies, such as carbon to nitrogen ratios ($\text{C}_{\text{org}}/\text{N}_{\text{tot}}$),
1219 bulk organic carbon isotopes ($\delta^{13}\text{C}_{\text{org}}$) or hydrogen and oxygen indices (HI, OI) from Rock-Eval
1220 pyrolysis, are useful tools to explore aspects of environmental variability. Although bulk proxies provide
1221 limited environmental information, they do have the potential to indicate major changes in sources
1222 and/or fluxes of organic matter (see Meyers, 2003 for a review). Furthermore, bulk analyses facilitate the
1223 formulation of detailed hypotheses and the design of targeted, high-resolution biomarker studies. A
1224 popular strategy for paleoenvironmental analyses therefore is to compare biomarkers with proxy data
1225 from palynology and bulk organic geochemistry. Examples include studies of East African lakes,
1226 combining pollen data with compound-specific carbon and hydrogen isotope data, lignin composition,
1227 and lake surface temperature data (Tierney et al., 2010; Berke et al., 2012).

1228 A basic problem is that biomarkers, as any organic substance left behind by a deceased organism, are
1229 affected by microbial and physicochemical degradation, which can occur before or after deposition in the
1230 sediments. Nitrogen- and oxygen-containing compounds, such as carbohydrates or amino acids, are
1231 prone to microbial degradation (see also section 2.2.3.), whereas molecules based on hydrocarbons, such
1232 as lipids, may preserve source-specific information over geological timescales, i.e., over hundreds of
1233 millions of years (e.g., Marynowski et al., 2011; Izart et al., 2012; Rohrsen et al., 2013). The oldest
1234 deep-time context from which biomarkers have been interpreted dates back ~1.6 billion years
1235 (Pawlowska et al., 2013). Microbial degradation may selectively alter the relative amounts of lipid
1236 biomarkers of different recalcitrance or that are bio-accessible, e.g., in a clay mineral matrix. In order to

1237 minimize diagenetic bias, biomarker proxies are generally based on ratios or distributions of compounds
1238 of the same compound class that also likely derive from similar source tissue types.

1239 In the following, we discuss some practicalities of biomarker approaches, focusing on the recovery of
1240 paleoenvironmental records from ancient lakes, and including strategic considerations in order to gain
1241 maximum and reliable information from biomarker studies.

1242 Lipid biomarkers are principally extracted from sediment samples by organic solvents through the
1243 application of a range of extraction techniques, in particular sonication, accelerated solvent extraction
1244 (ASE), and microwave-assisted solvent extraction (MAE), all of which are fast methods for high sample
1245 throughput (Camel, 2000; Kornilova and Rosell-Melé, 2003). Although time consuming and using
1246 greater quantities of solvent, Soxhlet extractions are a suitable option for the study of sediments with
1247 very low organic matter content. Once extracted, the lipids can be identified and quantified by gas
1248 chromatography-mass spectrometry (GC-MS). In cases where differences in ionization during GC-MS
1249 analyses may lead to quantitative bias, a standard gas chromatograph fitted with a flame ionization
1250 detector (GC-FID) is used for quantification instead. A typical application is the quantification of
1251 haptophyte-derived alkenones, which can provide information on paleo-surface water temperatures (Uk₃₇
1252 LST proxy, Table 3). Larger molecules, such as bacterial bacteriohopanepolyols (BHPs) or archaeal
1253 glycerol dialkyl glycerol tetraethers (GDGTs), which also provide temperature proxies for lake surface
1254 waters as well as for soils, are analyzed by high-performance liquid chromatography-mass spectrometry
1255 (HPLC-MS). Lignin phenols used for vegetation reconstructions or terrestrial input can be analyzed
1256 efficiently by pyrolysis-gas chromatography-mass spectrometry (Py-GC-MS; e.g., Ishiwatari et al.,
1257 2006). For compound-specific stable isotope analyses ($\delta^{13}\text{C}$, $\delta^2\text{H}$, $\delta^{15}\text{N}$), compound classes or individual
1258 compounds can be isolated either through standard flash chromatography or automated preparative
1259 methods (prepGC, prepHPLC).

1260 However, a single lab is rarely able to produce the complete suite of biomarker proxy data. One
1261 possibility to increase material- and cost-efficiency of multiproxy biomarker studies is to design a
1262 coordinated approach and to distribute splits of the total lipid extracts (TLEs) for various applications
1263 rather than bulk sediments samples.

1264 The inter- and intra-habitat heterogeneity of biomarkers leads to the fundamental problem that
1265 biomarker-based proxies often are not interchangeable between investigated lake basins. Therefore, it is
1266 highly recommended to include a survey of the modern biomarker sources within the catchment of an
1267 environmental archive (biogeochemical fingerprinting) when planning biomarker-based
1268 paleoenvironmental research. In this context, it is crucial to correctly identify the major sources of
1269 sedimentary organic matter. For example, rather than the living vegetation itself, plant litter is a major
1270 source of terrigenous organic matter. This is important as the biomarker composition of plant litter is
1271 already altered compared to living biomass due to degradation processes that begin as soon as a plant
1272 dies. Soils are another major organic matter pool within many lacustrine basins. Furthermore, as they
1273 accumulate over time, soils integrate the biogeochemical signature of the changing vegetation cover and
1274 its degradation products and of belowground biomass (root material plus fungal and bacterial biomass).
1275 Thus, biogeochemical fingerprinting of the major organic matter pools facilitates the correct
1276 identification of the sources of the organic matter in lake sediments and helps assessing potential input of
1277 pre-aged material, which can lead to considerable bias in high-resolution records (Douglas et al., 2014).
1278 A large uncertainty also affects the interpretation of stable isotope data, in particular, for hydrogen
1279 isotopes (Sachse et al., 2012). Determination of the carbon and hydrogen isotope composition of
1280 biomarkers from the major organic matter pools in the modern environment of a lacustrine catchment is
1281 needed in order to improve the understanding of the impact of hydrological changes on compound-
1282 specific isotope compositions (see Wilkie et al., 2013 for the Lake El'gygytgyn drilling project). For
1283 paleotemperature reconstructions, a surface sediment-surface water calibration is highly recommended
1284 due to the many factors that can bias biomarker-based temperature proxies in lacustrine settings.

1285 Unless sediment cores are stored in a freezer, microbial breakdown processes continue to alter
1286 organic matter in the sediments. Samples for biomarker studies should therefore be taken from the cores
1287 at the earliest possible occasion and then be frozen (ideally at $-80\text{ }^{\circ}\text{C}$) or freeze-dried as soon as possible
1288 to prevent further microbial degradation. Even after freeze-drying, cold storage is advisable in order to
1289 preserve labile biomarkers, such as mono- and poly-unsaturated fatty acids.

1290 During sampling and storage, it is crucial to avoid contamination with organic compounds derived
1291 from petroleum products such lubricants and plastics (also see section 2.1.2.). Polyethylene (PE), which

1292 is commonly used for soft plastics, such as sampling bags or the lids of sample containers, releases a
1293 series of compounds, in particular branched alkenes with quaternary carbon atoms (BAQCs), which can
1294 contaminate even freeze-dried sediment during prolonged storage (Brocks et al., 2008; Holtvoeth,
1295 unpublished data). Samples should be transferred into furnaced glass jars or vials, or into containers lined
1296 with combusted foil (aluminum foil heated to 450°C for at least 4 h).

1297 The study of biomarkers preserved in lacustrine sediments has developed strongly ever since their
1298 potential for paleo-environmental reconstructions was recognized and developed from the late 1960s
1299 onwards (e.g., Cranwell, 1973; Brooks et al., 1976). This is largely due to the rapid advance of analytical
1300 technology and the increase in analytical capacity. Promising analytical methods are established using
1301 advanced detectors for mass spectrometry, such as a quadrupole time-of-flight mass detector (GC/Q-
1302 TOF) or orbitraps, increasing resolution and precision in the detection of molecular fragments.
1303 Moreover, high-temperature GC-MS applications extend the range of GC-amendable compounds
1304 towards high-boiling biomarkers (e.g., wax esters, triacylglycerides, GDGTs; Sutton and Rowland,
1305 2012). A dynamic field with great potential for biogeochemical fingerprinting of archaeal and bacterial
1306 organic matter sources and reconstructions of microbial ecology is the analysis of compounds derived
1307 from bacteriohopanepolyols (BHPs; Talbot et al., 2003; Zarzycki and Portka, 2015) and of intact polar
1308 lipids (IPLs; Rethemeyer et al., 2010; Tierney et al., 2012; Buckles et al., 2014). A recent and entirely
1309 different approach to the use of biomarkers as indicators of ecosystem change is the application of
1310 advanced statistical programs integrated in the analytical software for identifying the key variables in
1311 biomarker screening data that may include hundreds of compounds in an environmental sample
1312 (lipidomics, environmental metabolomics; Bundy et al., 2009). Finally, genetic research targeting
1313 enzymatically controlled molecular adaptations of organisms to certain environmental conditions is
1314 fundamentally changing the conventional interpretation of biomarker data. Rather than linking a
1315 biomarker to the presence of a certain organism or groups of organisms, it can instead be assigned to a
1316 specific ecological niche (e.g., Welander et al., 2012). This change in view is important for the
1317 understanding of the geochemical fingerprint that past ecosystems left behind and opens up new
1318 possibilities to interpret assemblages of biomarkers that, individually, had been regarded as unspecific.

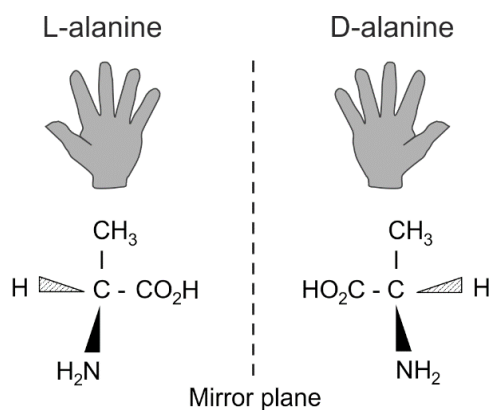
1319 2.2.5. *Ancient amino acids*

1320 Towards the end of the 19th century, the first identification of the likely remains of proteins in fossils
1321 paved the way for their use in the earth sciences. Amino acids, the building blocks of proteins, are found
1322 in all living tissues and can be preserved in subfossil biominerals, such as shells, as well as in sediment.
1323 Analysis of these ancient amino acids has proved important for three main avenues of research: dating
1324 (via amino acid geochronology), species identification (via paleoproteomics), and to assess the integrity
1325 of the organic matter for other biomolecular studies (e.g., aDNA).

1326 Amino acids can be used for dating purposes because of amino acid racemization (AAR), i.e., the
1327 time-dependent breakdown of proteins (and their constituent amino acids) in fossils. Spanning an age
1328 range from 10 years ago up to as long ago as a few million years, the method is applicable to the whole
1329 of the Quaternary Period (see Lowe and Walker, 2015 for a recent review). Advances in
1330 chromatography, preparation methods, and the choice of material for dating have greatly improved the
1331 accuracy of the methods, and demonstrate the technique's potential for developing regional Quaternary
1332 chronologies around the world (e.g., Penkman et al., 2011; Wehmiller, 2012). Identification of
1333 endogenous amino acids in Cretaceous and Paleogene samples (Miller and Hare, 1980; Penkman et al.,
1334 2013) opens up the opportunity to use other protein degradation reactions to date material over much
1335 longer timescales.

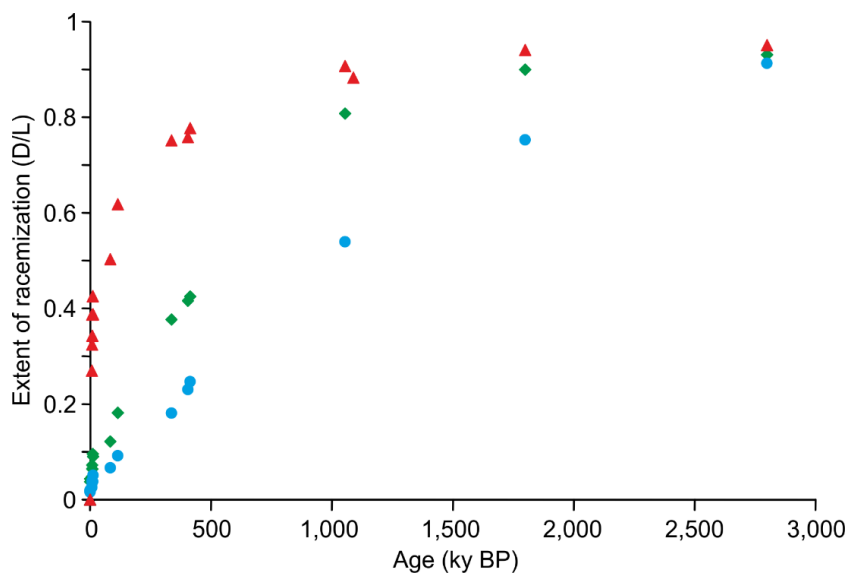
1336 The 20 naturally-occurring amino acids all have a central carbon atom (the α -C) with four attached
1337 groups: an amino group (NH₃), a carboxylic acid group (COOH), hydrogen (H), and a side chain (R) that
1338 defines the type of amino acid. In glycine, the side chain is H, but for all other amino acids, the α -C has
1339 four different groups (Fig. 9). The four distinct groups connected by single bonds make the α -C a chiral
1340 center, meaning that it can exist as two stereoisomers: the *laevo* (L-form) and *dextro* (D-form), named
1341 after the optical activity of glyceraldehyde. In living organisms, proteins are almost exclusively made
1342 from the L-form. However, this dominance of one form is thermodynamically unstable, so after death, a
1343 spontaneous reaction occurs to balance the abundance of both forms. The extent of AAR is analyzed by
1344 gas or liquid chromatography and recorded as a D/L value. AAR continues until a dynamic equilibrium
1345 is reached (usually D/L=1). First applied to fossil shells (Hare and Abelson, 1968), AAR geochronology

1346 measures the extent of this degradation in fossils as an index of relative age (an aminostratigraphy),
 1347 which can provide calibrated ages in combination with known-age samples or detailed temperature
 1348 records (Fig. 10). This then may allow correlation of deposits with the marine oxygen isotope stage
 1349 (MIS) record (for lacustrine deposits see McCoy, 1987; Bowen et al., 1989; Magee et al., 1995; Oviatt et
 1350 al., 1999; Kaufman, 2003a; Ortiz et al., 2004; Penkman et al., 2011), to a sub-MIS level for at least the
 1351 Late Pleistocene.
 1352



1353

1354 **Fig. 9.** Most amino acids have no plane of symmetry, just like hands, so their mirror images are non-
 1355 superimposable and therefore distinct from each other. The breakdown of left-handed molecules to the right-handed
 1356 form over time provides a mechanism for estimating age of fossil material.
 1357



1358

1359 **Fig. 10.** The increase in racemization in the opercula of the snail species *Bithynia tentaculata* with age for the free
 1360 amino acid (FAA) aspartic acid (Asx; red symbols) and the total hydrolysable amino acids (THAA) valine (Val;

1361 blue symbols) and alanine (Ala; green symbols) from British deposits with independent geochronology. Asx
1362 racemizes rapidly and is therefore most valuable for separating sites younger than MIS 7 in these temperate
1363 deposits. Val, in contrast, racemizes more slowly and is able to differentiate between sites back to the Pliocene, but
1364 provides poorer resolution for young sites. Utilizing multiple amino acids with different rates of degradation
1365 therefore enables greater time depth and age resolution. Image modified from Penkman et al. (2011).

1366
1367 Protein degradation consists of a series of chemical reactions that are dependent on time, but also on
1368 environmental factors (e.g., pH, availability of water, temperature), which can confound the time signal.
1369 These difficulties in AAR's early applications have led to a focus on analyzing 'closed-system' protein
1370 from fossil samples (Towe, 1980), where the fraction of protein analyzed is physically or chemically
1371 shielded from the environment. The chemically-isolated 'intra-crystalline' fraction found in mollusk and
1372 egg shells forms such a closed system, meaning that the AAR within this fraction is solely dependent on
1373 time and temperature, and therefore predictable (Brooks et al., 1990; Sykes et al., 1995; Penkman et al.,
1374 2008). AAR has been particularly successful in dating carbonate fossils (shells, eggshells, foraminifera,
1375 ostracods) and in long-lived biominerals (e.g., corals), providing age information *within* an individual
1376 sample (Hendy et al., 2012). In subfossil samples, the different proteins break down at different rates, so
1377 analyses are undertaken on monospecific samples (usually individual mollusk shells, a few mg in
1378 weight). Labs performing AAR have developed dating frameworks for a large number of commonly
1379 occurring species, but tests can be undertaken on additional species to examine whether they would be
1380 suitable for AAR dating. The crystal phase of calcite (e.g., opercula, eggshell, ostracods) are more stable
1381 over longer timescales and are therefore preferred for material of Early and Middle Pleistocene age
1382 (Penkman et al., 2011).

1383 The rate of breakdown towards D/L equilibrium in the intra-crystalline fraction is still affected by
1384 temperature, so comparative frameworks need to be applied from regions with a broadly similar
1385 temperature history. However where age control is available, the extent of racemization can then be used
1386 to estimate the effective diagenetic temperatures (Kaufman, 2003b). Published amino acid data are now
1387 being archived by NOAA and are freely available at <http://www.ncdc.noaa.gov/paleo/aar.html>.

1388 The advent of soft-ionization mass spectrometers made protein sequence identification more routine,
1389 and this was soon applied to fossil material (Ostrom et al., 2000). The ordering of the amino acids in a

1390 peptide chain (its sequence) can be diagnostic of the species from which it came, and therefore
1391 identification of specific ancient proteins informs on the past biota (Buckley and Wadsworth, 2014).
1392 While the evolutionary picture from proteins is not as detailed as that from aDNA (as changes in the
1393 peptide sequences are significantly slower), proteins are significantly more stable than DNA (see section
1394 2.2.6.), allowing identification of peptide sequences (and hence phylogenetic information) from material
1395 where aDNA is not recoverable (Cappellini et al., 2011). Focusing initially on Pleistocene bones and
1396 shell, this technique has primarily been applied to terrestrial deposits, but the excellent preservation of
1397 organic material in lake sediments (often due to anoxic conditions) lends itself to the expansion of
1398 paleoproteomics to lacustrine material.

1399 For all biomolecular analyses, it is critical to identify and exclude contamination. One advantage of
1400 protein analysis is that (unlike aDNA) proteins do not have to be amplified for detection, while the
1401 predictable nature of protein degradation enables identification of modern contamination (e.g., Buckley
1402 et al., 2008). Amino acid analysis of sedimentary material provided a useful tool for identifying the
1403 presence of original biomolecules, and hence helped authenticate the oldest DNA sequences yet
1404 recovered (Willerslev et al., 2007).

1405 A non-specialist can collect material and/or sediment samples in the field, and samples should be
1406 stored at or below room temperature. Biominerals for AAR dating are typically obtained from wet-
1407 sieved residues of sediment samples. Development of better preparative and analytical methodologies is
1408 allowing analysis of smaller and more degraded samples, while protein databases to match recovered
1409 sequences against are ever-growing, enabling more accurate species identifications. Amino acids and
1410 proteins are present in geological samples, and our abilities to use the geochronological and biological
1411 information they hold are advancing rapidly.

1412 2.2.6. *Ancient DNA*

1413 The analysis of ancient DNA (aDNA), i.e., DNA of long-dead specimens (Krause, 2010), has become
1414 an emerging field in evolutionary biology and paleoecology during the last decades (e.g., Hofreiter et al.,
1415 2001; Pääbo et al., 2004; Willerslev et al., 2014; Hagelberg et al., 2015; Birks and Birks, 2016). Ancient
1416 DNA provides a unique opportunity to assess paleo-biodiversity, and to unravel past evolutionary and

1417 environmental processes by comparing genetic information of fossil and extant organisms. Moreover, as
1418 the nucleotide sequence of DNA fragments ('DNA sequences') can be digitally encoded in a
1419 standardized way and stored in public databases, newly generated DNA sequences are directly
1420 comparable to previously generated data.

1421 Ancient DNA may be isolated from (parts of) specimens preserved in sediment cores (incl.
1422 mummified tissues, bones, teeth, or other biomineralized material) or, more commonly, as 'sedimentary
1423 aDNA', i.e., organismal DNA that has been released into the water or directly into the sediment, often in
1424 the absence of visible fossils (Parducci et al., 2013, 2015; Pansu et al., 2015). However, aDNA is often
1425 of poor quality compared to DNA from samples of extant organisms as DNA molecules exponentially
1426 degrade in smaller fragments after the death of an organism (Hofreiter et al., 2001; Allentoft et al., 2012).
1427 In addition, microorganisms may digest aDNA and even introduce mutations (Hofreiter et al., 2001;
1428 Krause, 2010 and Allentoft et al., 2012). These microbial effects are of particular concern for
1429 sedimentary aDNA analyses, as the respective DNA molecules are largely unprotected against the
1430 environment. Other factors determining degradation are, among others, time, environmental temperature,
1431 oxygen content, pH, sediment type and composition, and presence of free water (Allentoft et al., 2012;
1432 Hagelberg et al., 2015). Once a critical sequence length of approximately 15–25 nucleotides is reached,
1433 these DNA fragments lose their unique fingerprint and can no longer be used for comparative studies.

1434 The two main methods to decode aDNA are PCR-directed sequencing and direct sequencing. For
1435 PCR-directed sequencing, short oligonucleotides ('primers') have to be designed that selectively bind to
1436 a complementary aDNA target region. The advantage of this method is that a specific region (e.g., a
1437 particular gene or part of a gene) can be targeted, that the number of read errors is relatively low, and that
1438 contamination (e.g., through microbial DNA) is limited by the use of taxon-specific primers.

1439 Disadvantages are that the aDNA fragments targeted have to be relatively long (typically several
1440 hundreds of nucleotides) and the need to design specific primers, which is often difficult in the absence
1441 of reference data. In contrast, direct sequencing does not require specific primers and thousands or
1442 millions of short aDNA fragments can be directly encoded using 'next generation sequencing' platforms
1443 (e.g., Metzker, 2010; Mardis, 2011). These high-throughput methods work well with highly degraded
1444 aDNA fragments (i.e., with fragments as short as 25 nucleotides; Storvall et al., 2013), but the numerous

1445 individual aDNA reads have to be assembled using reference databases, which so far cover mainly
1446 ‘model’ organisms. Moreover, as no taxon-specific primers are used, contamination with exogenous
1447 DNA may constitute a major problem even though parts of these contaminations can be recognized *a*
1448 *posteriori* using bioinformatic tools (e.g., Schmieder and Edwards, 2011).

1449 Despite numerous methodological advances, aDNA analyses from lake-core materials remain
1450 problematic. Reasons are the small amount of available material, the highly degraded nature of aDNA
1451 from sediment cores, read errors, contamination, and incomplete reference databases (*sensu* Krause,
1452 2010; Pedersen et al., 2013; Birks and Birks, 2016). Therefore, aDNA from lake sediments is currently
1453 mainly used for biodiversity assessments of Late Pleistocene and Holocene communities (Bissett et al.,
1454 2005; Anderson-Carpenter et al., 2011; Boessenkool et al., 2014; Pansu et al., 2015), particularly as
1455 complementary information to fossil data (*sensu* Jørgensen et al., 2012; Parducci et al., 2013).

1456 Though respective analyses from ancient lake cores are still lagging, we expect to see more paleo-
1457 biodiversity assessments based on aDNA analyses in future deep drilling projects. Given that shallow
1458 areas in ancient lakes are typically more biodiverse than deep sites, we suggest to retrieve aDNA samples
1459 from near-shore sediment cores (although temperature and O₂ content in these areas might be higher,
1460 thus accelerating degradation; see also sections 2.1.1. and 2.2.1. about the disadvantages of near-shore
1461 sites and Cohen, 2012 for issues related to geological conditions in general). A principle problem is
1462 contamination with exogenous DNA. Therefore, cores have to be sampled under aseptic conditions and
1463 sedimentary aDNA sampling is ideally done directly in the field. However, previous studies have shown
1464 that aDNA can also be isolated from lake sediments after long-term refrigeration at 4°C (Bissett et al.,
1465 2005). Respective procedures of field sampling are similar to those used in deep biosphere studies
1466 (Ariztegui et al., 2015; also see section 2.2.4.). For this reason, coupling aDNA and subsurface biosphere
1467 studies in ancient lake drilling projects is advisable.

1468 Future studies will probably use direct high-throughput sequencing, perhaps extending the time frame
1469 of analyses to the Middle Pleistocene. However, the usefulness of high-throughput approaches may also
1470 increase once more extensive reference DNA datasets are established that include genetic information on
1471 both fossil and extant species. We therefore recommend to couple future aDNA studies in ancient lake
1472 drilling projects with genetic and genomic studies on extant species (see also section 3.3.).

1473 2.3. *Data accessibility and storage*

1474 Long-term core and sample storage as well as data accessibility and curation are of fundamental
1475 importance in light of the immense financial, logistic, labor, and intellectual efforts associated with
1476 scientific drilling projects. Over the last decades, core repositories (e.g., Bremen Core Repository
1477 (BCR), Bremen, Germany; Kochi Core Center (KCC), Kochi, Japan; National Lacustrine Core
1478 Facility (LacCore), Minneapolis, USA), equipped for core processing and particularly designed for the
1479 long-term storage of cores, samples, and data were established in order to service the scientific drilling
1480 community and to guarantee the accessibility of samples, relevant drilling metadata, and initial core
1481 descriptions (ICD) for extended periods. Project parties of scientific drilling projects are
1482 conventionally obliged to generate ICD-data, which typically encompasses whole core and split core
1483 MSCL data, split core surface imaging, visual core descriptions, smear slide analysis, and core
1484 correlation points (Fig. 3). ICD data is crucial for subsequent sample selection by those directly
1485 involved, or others that intend to study samples after the moratorium has ended (commonly two years
1486 after drilling).

1487 All relevant metadata and initial datasets generated on-site during drilling and laboratory-based
1488 core processing should, by default, be uploaded into dedicated, specific databases such as the ICDP
1489 Drilling Information System (DIS) for archival in professionally managed server facilities. In order to
1490 assure globally unique identifiers (Sample ID's) for long-term traceability between samples and data,
1491 International Geo Sample Numbers (IGSNs; <http://www.geosamples.org/aboutigns>) for each type of
1492 sample generated should be registered and assigned appropriately through the System for Earth
1493 Sample Registration (SESAR, <http://www.geosamples.org/>) by each project. Furthermore,
1494 fundamental datasets along with datasets published in scientific publications by scientific drilling
1495 project parties have to be made available through online databases such as PANGAEA
1496 (<http://www.pangaea.de/>) and/or NOAA National Geophysical Data Center Index to Marine and
1497 Lacustrine Geological Samples (<http://www.ngdc.noaa.gov/mgg/curator/curator.html>) in accordance
1498 with ICDP and national funding agency rules.

1499 However, most of the data storage and data sharing policies outlined above only apply to primary
1500 data generated from sediment cores. In contrast, there are still no generally accepted rules for storing
1501 and sharing secondary data and materials obtained during drilling campaigns. As a minimum,
1502 geological and biological voucher materials should be deposited in a freely accessible and
1503 internationally recognized scientific collection. Moreover, data should be made available through
1504 major public databases such as NCBI's GenBank (<http://www.ncbi.nlm.nih.gov/genbank>), the
1505 Paleobiology Database (<https://www.paleobiodb.org>), or the citable Dryad Digital Repository
1506 (<https://datadryad.org>).

1507 **3. Integrating geological and biological data**

1508 In section 2, we have shown that many different types of data, both geological and biological, can
1509 be obtained from drilling campaigns. Despite this rich variety of data types, lake drilling long
1510 remained the domain of earth scientists. As a result, many geological and paleolimnological analyses
1511 are well established and have been reviewed abundantly before (e.g., Cohen, 2003; O'Sullivan, 2004).
1512 In comparison, the use of organismal approaches to sediment-core data for questions related to
1513 environmental and evolutionary biology are newer, and therefore we focus here on data analysis
1514 methods that fall within this scope. We start with paleolimnological reconstructions as a nexus of
1515 environmental inquiry between strictly geological and biological approaches, then discuss the
1516 consequences of environmental change for biological diversification processes, and finally deal with
1517 aspects of the natural history of biota, i.e., timing evolutionary events and tracing character evolution.

1518 *3.1. Paleolimnological reconstruction*

1519 A major goal of scientific drilling projects in ancient lakes is the reconstruction of their
1520 paleolimnology and paleoclimate (Cohen, 2012). Sediment records revealed, for example, regional
1521 climatic and environmental expressions of Milankovitch/glacial-interglacial cycles (e.g., Hooghiemstra,
1522 1989; Kashiwaya et al., 2001; Wagner et al., 2014a) and their extremes in polar and tropical realms (e.g.,
1523 Cohen et al., 2007; Melles et al., 2012).

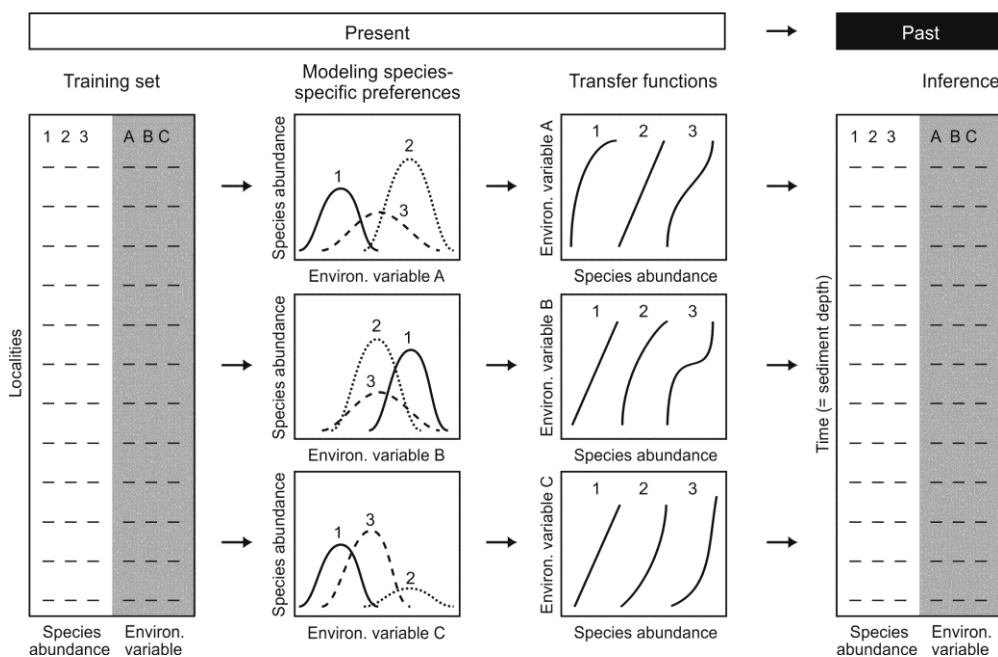
1524 Traditionally, paleolimnological reconstructions are primarily based on micro- and microfossils, both
1525 from sediment cores (primary data) and outcrops (secondary data) (see sections 2.2.1., 2.2.2.). However,
1526 modern reconstructions also integrate sedimentological, geochemical, and isotope information, thus
1527 enabling a more comprehensive assessment of paleoconditions in a given lake and its watershed. These
1528 conditions include, for example, past lake-level fluctuations (Filippov and Riedel, 2009) or changes in
1529 water depth (Lyons et al., 2016; Magyar et al., 2006), oxygen conditions (Costa et al., 2015), salinity
1530 (Mischke et al., 2010), water temperature (Goodwin et al., 2003; Castañeda and Schouten, 2011;
1531 Recasens et al., 2015), or primary productivity (Langlet et al., 2007; Recasens et al., 2015).
1532 Paleolimnological reconstructions often provide a well-informed picture about past regional (e.g.,
1533 Bergner and Trauth, 2004; Stager et al., 2009) or local environmental conditions (e.g., Mourguiart et al.,
1534 1998). Although the resolution and the indicative power of paleolimnological reconstructions largely
1535 depends on the availability and quality of the respective fossil record, preserved biological recorders
1536 sometimes reflect past climate and environmental changes even on the scale of decades or years (e.g.,
1537 Ghinassi et al., 2015).

1538 Analytical approaches for paleolimnological reconstructions range from qualitative to semi-
1539 quantitative and quantitative analyses. Qualitative inference is based on linking ecological and
1540 environmental information using indicator species, whereas semi-quantitative analyses utilize estimated
1541 changes in fossils proportions and involve a link between environmental and biotic information (Cohen,
1542 2003; Birks et al., 2012; also see section 2.2.). These approaches usually employ models
1543 ('paleoecological transfer functions', Fig. 11) that correlate modern species assemblages (i.e.,
1544 composition or abundances) and their environmental characteristics ('calibration'; Cohen, 2003; Juggins
1545 and Birks, 2012). Subsequently, fossils are used to infer paleoenvironmental conditions based on the
1546 revealed assemblage-environment relationship. Transfer functions require a reference data set (=
1547 'training set'), i.e., a data matrix that contains information on the distribution and abundance of recent
1548 species, either found in the water column or in surface sediments, together with information on the
1549 environmental variables that drive these distributional patterns. Ideally, a training dataset should be
1550 produced under controlled laboratory conditions, subjecting an organism or community (e.g.,
1551 phytoplankton) to a range of environmental conditions for establishing a causal link between the

1552 environmental (independent) variable and a set of response (dependent) variables (Saros et al., 2012).
 1553 However, because of the often complex relationship between environmental and distributional data,
 1554 these models are typically built with data from a range of modern lakes.

1555 For ancient lakes, the training of transfer functions might be restricted to the lake itself due to the lag
 1556 of analogues for endemic taxa (see also section 2.2.2.; Mackay et al., 2003). However, care is required,
 1557 because the non-independency of the training and testing datasets may cause statistical problems (for
 1558 details see Cohen, 2003).

1559



1560

1561

1562 **Fig. 11.** Flow of transfer function development for paleolimnological reconstructions. Species abundances (Arab
 1563 numerals) and environmental variables (letters) are measured in extant assemblages. These data are used to model
 1564 species-specific curves of environmental preferences, which then can be utilized to develop and refine predictive
 1565 functions of environment-abundance relationships using fossil assemblages from time-resolved sediment-core
 1566 samples in the lake of interest (modified from Fritz et al., 1999).

1567

1568 Whereas it is typically straightforward to produce environmental and biological datasets from extant
 1569 lakes, it is naturally more difficult to obtain the necessary fossil record from sediment cores. Of concern
 1570 are site selection (i.e., drilling shallow or deep sites; see also section 2.1.1.), continuity of records, as well
 1571 as data quality and resolution. The latter aspects are particularly important because they strongly affect
 1572 the applicability of transfer functions. It is also important to account for diagenetic and taphonomic

1573 processes that might mask or even mislead interpretations (see sections 2.2.1., 2.2.2.). Environmental
1574 reconstruction based on transfer functions and sediment proxies can give conflicting information for the
1575 interpretation of the same paleoenvironment (Cohen, 2003). Therefore, it is necessary to assure that there
1576 is, indeed, a causal link between the physico-chemical variable and the bioindicators' response (Juggins,
1577 2013). Strategies have to be developed to implement uncertainty in paleolimnological reconstructions.
1578 Another challenge is the need to disentangle individual and joined effects of multiple causal factors in
1579 the species-environment relationships (Juggins, 2013). Moreover, the species-environment relationship
1580 needs to be constant over time (i.e., no niche shift; for a critical assessment of transfer functions see also
1581 Juggins, 2013). These environmental variables should cover the range of environmental variation in the
1582 particular time frame of interest.

1583 Systematic knowledge and taxonomic concepts, in general, are major sources of misinterpretation of
1584 communities, meta-communities, and related ecological measures. For example, the biological relevance
1585 of morphological characters for species delimitation in the fossil record (e.g., diatoms, ostracods,
1586 mollusks) should be studied in extant analogs ideally using an integrative taxonomic approach based on
1587 phylogenetic relationships and character evolution analyses (see section 3.4.). New methods to be
1588 developed should also identify and estimate the range of effects of diagenetic and taphonomic processes
1589 on the records used for paleolimnological reconstructions (Birks et al., 2012).

1590 Undoubtedly, we will see several technical and methodological advances in the future, including
1591 improved dating of sediment records (e.g., Shanahan et al., 2013; Zanchetta et al., 2015; see also section
1592 2.1.5.), high-resolution datasets (Lacey et al., 2015), and high-throughput technologies (Tolu et al.,
1593 2015), allowing a more efficient production of long-term datasets that can be used for paleolimnological
1594 reconstructions (also see Seddon et al., 2014).

1595 The application of new proxies for paleolimnological reconstructions, such as biomarkers, is rapidly
1596 increasing as the relative costs for analytical instrumentation are coming down (section 2.2.4.). By now,
1597 a broad range of proxies has been established for lacustrine sediment records. These markers allow
1598 tracing temperatures, both surface water and mean annual air temperature, and moisture supply as key
1599 climatic factors in habitat dynamics (reviewed in Castañeda and Schouten, 2011). Paleothermometry
1600 based on estimations of racemization rates of different amino acids is another direction of analytical

1601 developments (Collins and Demarchi, 2014; see also section 2.2.6.). Further aspects of ecosystem change
1602 that can be targeted through biomarker proxy applications are photic zone anoxia or the dynamics of the
1603 terrestrial surroundings, including vegetation change, biomass burning, and soil erosion (see section
1604 2.2.4), all of which affect the trophic state of a lacustrine ecosystem.

1605 Despite the exiting results derived from integrating geological and biological data for
1606 paleolimnological reconstructions, there are many questions remaining how past ecosystem dynamics
1607 have shaped extant structures, resilience, and dynamics of ancient lakes. Future drilling projects might
1608 also aim at identifying early ecosystem warning signals for catastrophic regime shifts, and they may
1609 attempt to identify common environmental signals in multiple records spanning different spatial and
1610 temporal scales. Methods should be implemented that address problems of time lead and lag, conflicts
1611 between datasets derived from the various proxies used, data uncertainties, and spatial autocorrelation.

1612 3.2. *Consequences of environmental change: A biological diversification perspective*

1613 The often extraordinary endemic species richness and high morphological disparity of ancient lake
1614 taxa have inspired evolutionary biologists to unravel the underlying mode, tempo, and drivers of
1615 biological diversification. Increasingly, non-biologists, such as geologists and climatologists, are also
1616 interested in the balance between speciation and extinction events – the two sides of biological
1617 diversification. Their reasoning is straightforward: linking environmental and evolutionary data in space
1618 and time may help inferring and possibly even quantifying the consequences of past geological and/or
1619 climatic change. Moreover, information on past speciation and extinction events can be used to assess if
1620 and how an ecosystem can buffer such perturbations, i.e., its ecosystem resilience. Finally, an
1621 understanding of how abiotic factors have driven diversification processes in the past might help to
1622 predict the impact of future environmental changes on the biotic world under various global change
1623 scenarios (Condamine et al., 2013; Lawing and Matzke, 2014).

1624 Indeed, integrating chronologically constrained primary geological and biological datasets from
1625 sediment cores has become a quasi-standard in interdisciplinary drilling projects. Increasingly, these data
1626 are supplemented with secondary data such as fossil or stable isotope information from outcrops (see

1627 sections 2.1.6., 2.2.1., 2.2.2.) or with genetic information from extant species (section 3.3.). Three main,
1628 non-exclusive objectives are of potential interest:

1629 i) Inferring the drivers of diversification; of concern are, for example, potential effects of changes in
1630 temperature, lake-level fluctuations, and tephra depositions on speciation and/or extinction events
1631 (e.g., Schultheiß et al., 2009; Brown et al., 2010; Wagner et al., 2014c; Jovanovska et al., 2016).

1632 ii) Inferring the tempo of speciation; of relevance are whether, for example, environmental perturbations
1633 affect changes in diversification rates over time and triggered radiation events (e.g., Day et al., 2008;
1634 Sherbakov, 1999; Schön and Martens, 2011; Wysocka et al., 2014).

1635 iii) Inferring the mode of speciation; of interest are, for example, whether intralacustrine barriers or lake-
1636 level associated vicariance events have promoted allopatric speciation (Trajanovski et al., 2010;
1637 Koblmüller et al., 2011; Schreiber et al., 2012) or whether environmental changes opened new
1638 ecological niches, thus promoting adaptive radiation (e.g., Schön and Martens, 2004; Herder et al.,
1639 2008; Young et al., 2009; Muschick et al., 2012, 2014).

1640 Previous studies, for example, have shown that lake-level changes in ancient lakes had very different
1641 effects on their endemic taxa. Depending on the magnitude of lake-level changes, the ecosystem
1642 resilience of the respective lake, and the resistance of its biota, these changes may have caused massive
1643 extinction events across higher taxa (e.g., Schultheiß et al., 2009, 2011), selective extinction in some but
1644 speciation in other taxa (e.g., Kroll et al., 2012; Takahashi and Moreno, 2015; Weiss et al., 2015), or no
1645 visible effect on patterns of diversification (e.g., Föller et al., 2015).

1646 However, linking geological and biological data from drilling projects, particularly in the context of
1647 hypothesis testing, is not always straightforward (see also section 3.4.). This issue is particularly true
1648 when secondary data are used, which are not *per se* chronologically constrained with the primary
1649 sediment core data. For example, linking a sequence of diversification events inferred from DNA
1650 phylogenies to a series of lake-phases (e.g., Kroll et al., 2012) is challenging because of uncertainties in
1651 the timing of both datasets. Furthermore, even when a correlation can be demonstrated, it does not imply
1652 causality. In addition, drilling projects largely focus on the generation of abiotic data, which may not
1653 affect tempo and mode of diversification directly. Instead, abiotic changes possibly act indirectly through
1654 biotic drivers such as changes in character states (Hansen, 2014; Cantalapiedra et al., 2014; Salzburger et

1655 al., 2014; see also section 3.4.) or community structures (Hauffe et al., 2015). Moreover, to unravel the
1656 consequences of environmental change (*sensu* Condamine et al., 2013; Lawing and Matzke, 2014), it is
1657 not only important to understand if abiotic drivers are involved, but also to what extent. These problems,
1658 however, are not unique to scientific deep drilling projects (Rabosky and McCune, 2010).

1659 In the following, we therefore exemplify recent developments of statistical approaches that allow
1660 testing specific hypotheses about drivers of diversification. Depending on the type of dataset to be linked
1661 in drilling projects, three categories of analyses can be distinguished.

1662 The first set of methods (secondary + secondary datasets) is based on dated molecular phylogenies
1663 and tests for deviation from a constant diversification rate (see also section 3.3. and Morlon, 2014). In
1664 particular, the methods of Stadler (2011a, 2011b), Morlon et al. (2011), and Rabosky (2014) allow
1665 detecting distinct shifts and a non-linear time-dependence of diversification rates. These molecular data
1666 can be complemented by another set of secondary data, i.e., biotic factors hypothesized to drive
1667 diversification such as habitat use or changes in morphological character states (see also section 3.4. and
1668 the review of Ng and Smith, 2014).

1669 The second set of methods (secondary + primary datasets) uses dated molecular phylogenies in
1670 combination with abiotic factors that can be directly obtained from the sediment core and/or the
1671 borehole. Though the idea of abiotic control on diversification ('Court Jester hypothesis'; Barnosky,
1672 2001) is conclusive, to the best of our knowledge, only two approaches exist for testing the influence of
1673 abiotic factors on speciation and extinction events (Condamine et al., 2013; Cantalapiedra et al., 2014).
1674 The method of Cantalapiedra et al. (2014), for example, enables the identification of the relative
1675 importance of abiotic versus biotic drivers, their consequences on diversification rates, and the change of
1676 rates over time. Previous studies have indicated an interplay of present-day abiotic factors, such as area
1677 and depth of the African Great Lakes, and biotic characteristics, such as mating behavior, in determining
1678 the probability of cichlid lineages to form intralacustrine species flocks (Wagner et al., 2012).

1679 The third set of methods (primary + primary datasets) for inferring drivers of biological
1680 diversification uses dated origination and extinction events obtained from fossils, their biotic features,
1681 and abiotic conditions. The estimation of diversification rates based on fossil occurrence including
1682 taphonomic uncertainties is constantly improving (Foote, 2000), and the influence of abiotic drivers on

1683 these rates can be assessed using standard regression techniques (e.g., Neubauer et al., 2015). However,
1684 an alternative framework suggested by Silvestro et al. (2014) enables a simultaneous inference of
1685 diversification rates and identifies how these rates are shaped over time by biotic or abiotic factors. This
1686 allows including various factors such as morphological changes, standing species richness for diversity-
1687 dependent diversification, and environmental changes.

1688 Whereas these three sets of methods have largely improved our ability to test evolutionary hypotheses
1689 based on data derived from interdisciplinary deep drilling projects, they all share the same major
1690 limitations and pitfalls:

- 1691 i) Typically only the influence of a single abiotic or biotic driver of diversification can be estimated
1692 due to the current lack of tests for multiple drivers (Rabosky and McCune, 2010; Morlon, 2014).
- 1693 ii) Most analyses require a large set of species for correctly identifying the effect of a driver (Davis et
1694 al., 2013). Though many ancient lake species flocks are relatively species rich, the statistical power
1695 for smaller flocks might not be sufficient. For such cases, simulations have been suggested (Rabosky
1696 and Goldberg, 2015). Another option to increase the power is to test for a common signal across
1697 species groups by combining the phylogenetic information from several species flocks (i.e.,
1698 phylogenetic meta-analyses; Adams, 2008).
- 1699 iii) Constraining primary and secondary datasets from drilling projects is often challenging because they
1700 are independently dated. Moreover, the dating uncertainties of primary datasets from sediment cores
1701 (see section 2.1.5.) may be lower by orders of magnitudes than those of some secondary datasets
1702 (Wilke et al., 2009; see also section 3.3.).
- 1703 iv) Environmentally-triggered diversification events frequently occur with a time lag (Stadler, 2011b;
1704 Dynesius and Jansson, 2014), making a direct link of geological and biological datasets difficult. A
1705 solution for this problem could potentially come from the field of epidemiology where the spatial
1706 extent and duration of disease outbreaks and their causes are analyzed (e.g., Liang et al., 2010). In
1707 interdisciplinary drilling projects, similar approaches (e.g., marked point pattern analyses; Ripley,
1708 1976) could be used to identify abiotic drivers and the temporal extent of their influences on species
1709 diversification.
- 1710 v) The ability to estimate extinction rates based on phylogenies remains difficult (Rabosky, 2010) and

1711 the incorporation of the fossil record has been advocated (Quental and Marshall, 2010). However, as
1712 macrofossils are rare in sediment cores (section 2.2.1), statistical approaches would need to be
1713 applied that enable inferring extinction rates based on phylogenies alone (e.g., Morlon et al., 2011).
1714 vi) Another problem involves the suitability of the model taxa studied. Many ancient lakes harbor a high
1715 number of endemic species, often with diverse niches and different life styles (e.g., benthic vs.
1716 planktonic and generalist vs. specialist). However, baseline studies are necessary to verify that the
1717 candidate taxa, indeed, represent suitable model systems for the questions of interest in terms of, for
1718 example, monophyly and species richness (for details of sampling requirements and potential
1719 sampling biases see section 3.3.).

1720
1721 In general, the candidate abiotic and biotic drivers to be studied in interdisciplinary deep drilling
1722 projects have to be selected based on the specific scientific objectives. Apart from ‘standard’ abiotic
1723 factors such as lake-level changes, desiccation, and salinization events, parameters such as temperature
1724 and productivity (Condamine et al., 2013) have been proposed to influence diversification events and
1725 rates. Biotic drivers might be even more diverse. The underlying mechanisms and causality of some
1726 biotic drivers are well understood (e.g., the pharyngeal jaw of cichlids, Salzburger et al., 2014; depth
1727 preferences, Stelbrink et al., 2015), whereas the influence of species richness and composition on
1728 diversification is controversially discussed (e.g., Day et al., 2008; Harmon and Harrison, 2015; Hubert et
1729 al., 2015; Rabosky and Hurlbert, 2015). Also, the timeframe of evolution has to be considered when
1730 linking geological and biological datasets. Some ancient lakes such as lakes Baikal (e.g., Müller et al.,
1731 2001) and Ohrid (e.g., Albrecht and Wilke, 2008) are considered to harbor many old species or groups of
1732 species. Other lakes, though being old, experienced a series of major environmental perturbations, and
1733 the respective endemic species are often comparably young. Examples are Lake Malawi (e.g., Genner et
1734 al., 2007), the Caspian Sea (e.g., Dumont, 1998), and Lake Titicaca (e.g., Benavides, 2005; Kroll et al.,
1735 2012). Due to the young age of many taxa and problems of incomplete lineage sorting (for details see
1736 section 3.3.), phylogenetic relationships may be difficult to ascertain. In such cases, the use of primary
1737 data, such as fossils from the sediment core, is recommended for the study of old and/or long-term
1738 diversification processes, though sufficient data are rarely available.

1739 3.3. *Molecular dating*

1740 The ability to reconstruct older evolutionary events based on aDNA (see section 2.2.6.) remains
1741 limited. However, the lack of aDNA may be compensated for using DNA information from extant
1742 specimens. During organismal evolution, mutations become fixed over time (= substitutions) in the
1743 genome. Comparing these substitutions using DNA sequences of individuals that share a common
1744 ancestor may allow to reconstruct past evolutionary events, potentially driven by ecological,
1745 environmental, and/or geological processes (*sensu* Avise, 2000). A large ancestral population of
1746 lacustrine animals may, for example, become separated into two geographically separated
1747 subpopulations due to a severe lake-level lowstand. These resulting subpopulations might then evolve
1748 independently. After time, the two subpopulations thus become genetically distinct and potentially even
1749 new species. By comparing the substitutional patterns of the extant subpopulations/species using
1750 statistical approaches, the genetic structure of the common ancestor of these populations/species at the
1751 time of population subdivision (= ‘most recent common ancestor’) can be reconstructed. Under certain
1752 conditions, these divergence events might be time-calibrated using a methodology called molecular
1753 dating. This approach is based on the molecular clock hypothesis (Zuckerkandl and Pauling, 1965),
1754 which relates nucleotide or amino acid substitutions to time. In other words, knowing the number of
1755 substitutional differences between extant populations/species may allow for calculating divergence time
1756 and thus the age of the underlying event that potentially drove the separation of the ancestral population.

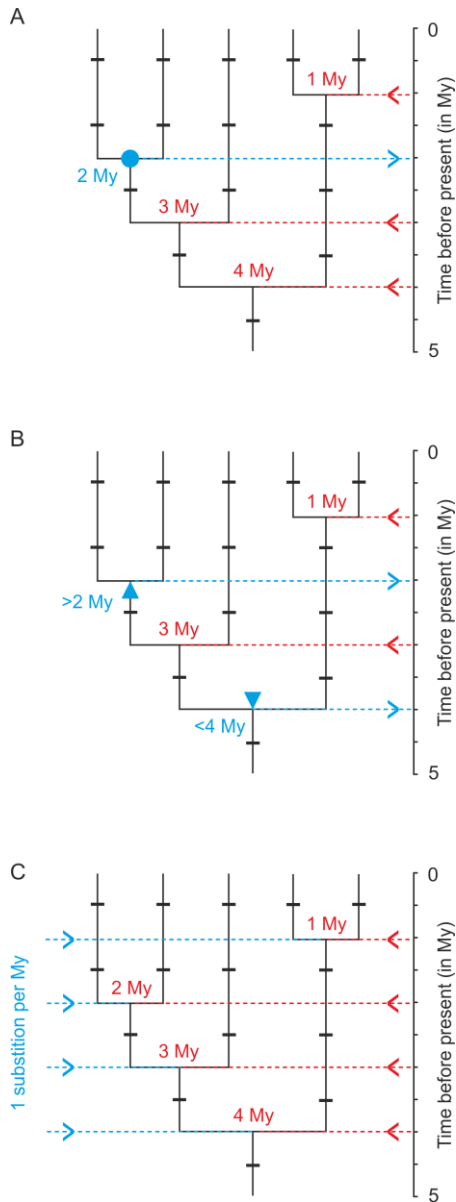
1757 Today, molecular dating is widely used in evolutionary biology to infer such past evolutionary
1758 events. However, the accuracy and even the applicability of molecular dating have long been subject to
1759 controversy (e.g., Takahata, 2007; Wilke et al., 2009; Ho, 2014; Ho and Duchêne, 2014). Substitution
1760 rates may, for example, vary among taxa, with life history traits, and/or over time, making a precise
1761 dating of evolutionary events challenging. They are gene-specific and highly sensitive to problems such
1762 as incomplete lineage sorting and substitutional saturation. Moreover, results from molecular dating
1763 analyses can easily be misinterpreted if, for example, the sampling design of extant species is insufficient
1764 (e.g., Wilke, 2004; Wilke et al., 2009; Friedman et al., 2013). For most of these problems, test statistics
1765 and mathematical solutions are available, which have made molecular dating statistically sound.

1766 However, a problem that continues to exist is that the molecular clock has to be calibrated in order to be
1767 able to calculate absolute times. This can be done using calibration points or bounds from externally
1768 derived dates such as ages of fossil occurrences and biogeographical events (e.g., vicariance events
1769 resulting from the closure of the Isthmus of Panama or the Mediterranean Salinity Crisis; Ho et al.,
1770 2015). Alternatively, calibration can be achieved via external clock rates that are gene- and often also
1771 taxon- or trait-specific (reviewed in Wilke et al., 2009). All of these different calibration methods have in
1772 common that uncertainties introduced by the calibration process often far outcompete mathematical
1773 problems with the clock approach.

1774 The two main approaches that are currently used for molecular dating are molecular clock and
1775 coalescence analyses. The former typically calculates divergence times between species or groups of
1776 species by estimating the number of substitutions that occurred along the respective branches of a
1777 phylogenetic tree (Fig. 12). The latter models the timing of demographic and spatial expansion events of
1778 populations (e.g., mismatch analyses; Rogers and Harpending, 1992), or past changes in population sizes
1779 (e.g., Bayesian Skyline Plots; Ho and Shapiro, 2011).

1780

1781



1782

1783

1784

Fig. 12. Methods for calibrating molecular clock trees. Substitutions along the branches of a phylogenetic tree are

1785

shown as black rectangles. A) Calibration with point(s) from externally derived dates. In this example, the known

1786

age of a divergence event (blue circle and blue dashes line) is used to estimate the timing of three other divergence

1787

events in the tree (red dashed lines). B) Calibration with bounds from externally derived dates. In this example, the

1788

known minimum and maximum ages of two divergence events (upward and downward pointing blue arrows,

1789

respectively, and blue dashes lines) are used to estimate the timing of two other divergence events in the tree (red

1790

dashed lines). C) Calibration with a known external molecular clock rate. In this example, a fixed external clock

1791

rate of 1 substitution per My and derived depth of nodes in the tree (blue dashed lines) are used to calculate absolute

1792

divergence ages (red dashed lines). For reasons of clarity, error bars of calibration uncertainties are not shown.

1793

1794 Both molecular clock and coalescence analyses have been used extensively in ancient lake studies
1795 including lakes Baikal (e.g., Sherbakov, 1999; Koskinen et al., 2002; Fazalova et al., 2010), Tanganyika
1796 (e.g., Nevado et al., 2013; Koblmüller et al., 2015), Malawi (e.g., Genner et al., 2010; Schultheiß et al.,
1797 2011), and Titicaca (e.g., Kroll et al., 2012). However, despite the high potential of molecular clock
1798 approaches, as yet there are only few examples for the application of molecular dating in ancient lake
1799 drilling projects. Wagner et al. (2014b), for example, found a temporal correspondence between a lake-
1800 level low stand in Lake Prespa on the Balkan as inferred from sediment-core data, the Late Pleistocene
1801 Toba eruption, and a spatial expansion in the lake's most abundant mussel species. Moreover, Föller et
1802 al. (2015) used lineages-through-time plots (i.e., a visualization of the number of accumulated
1803 evolutionary lineages over time inferred from molecular clock analyses; Harvey et al., 1994) and
1804 diversification-rate analyses in an attempt to link major environmental events inferred from the deep
1805 drilling project in Lake Ohrid to changes in diversification rates of endemic species.

1806 Certainly, molecular clock analyses will gain more importance in future deep drilling projects.
1807 However, given the 'vagaries' (Ayala, 1997) of the molecular clock, we recommend a careful planning
1808 of molecular dating approaches. This concerns the choice of methods, molecular markers, calibration
1809 means, and taxa. The latter requires especially thorough planning because sampling in ancient lakes can
1810 be expensive and time consuming, and incomplete sampling may bias molecular dating results. As a rule
1811 of thumb, molecular clock analyses in ancient lakes work best with large, monophyletic groups of
1812 endemic species – so-called species flocks (e.g., Schön and Martens, 2004). They typically evolved
1813 within the lake and thus are likely to reflect its environmental, ecological, and geological history.

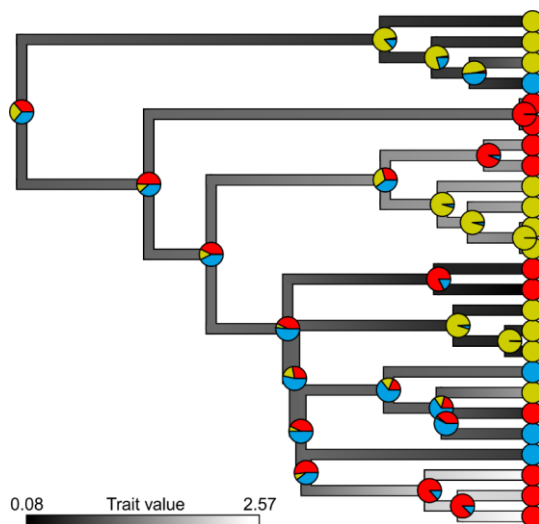
1814 Research on the behavior of the clock, calibration means, and data basis is continuing at high pace.
1815 Of interest for drilling projects in ancient lakes are, for example, newly developed models that enable
1816 improved fossil calibrations (e.g., Heath et al., 2014; Gavryushkina et al., 2015). However,
1817 notwithstanding the progress that will be made in the future, confidence intervals of molecular dating
1818 results will likely remain wide. It is therefore important to consider these uncertainties in all conclusions
1819 made based on molecular clock or coalescence approaches in order to avoid misinterpretations (reviewed
1820 in Hipsley and Müller, 2014; Warnock et al., 2015).

1821 3.4. *Tracing character evolution*

1822 Understanding character (= 'trait') evolution of a species' individual characters over time, such as
1823 changes in morphological/anatomical traits, ecological niches, functional roles in ecosystems,
1824 reproductive modes, or changes in geographic distributions, is of great relevance for linking geological
1825 and biological histories in ancient lakes. Tracing character evolution ideally involves groups of species
1826 that originated within the lake (species flocks) and which are comparatively old, thus allowing for
1827 inferring character evolution along the lake's entire geological history.

1828 Studies of character evolution can be done utilizing two main approaches. The first involves the
1829 direct observation of character-state change using chronological fossil information, often supplemented
1830 with other primary information obtained from the sediment cores. Examples include stratigraphic series
1831 of gastropods (Williamson, 1981; Van Bocxlaer and Hunt, 2013) and diatoms (e.g. Khursevich, 2006).
1832 However, continuous, high quality fossil information is typically not available (see section 2.2.1.) or only
1833 for selected groups of microfossils (section 2.2.2.). In the latter case, secondary data have to be used, i.e.,
1834 information on characters states of extant species together with a hypothesis about their evolutionary
1835 relationships (typically a phylogenetic tree).

1836 There are two particular interests in tracing character evolution along a phylogenetic tree. The first
1837 involves 'ancestral state reconstruction', i.e., the reconstruction of either categorical or continuous
1838 character states back in time (sometimes also referred to as 'character mapping' or 'character
1839 optimization'; see Fig. 13).



1840

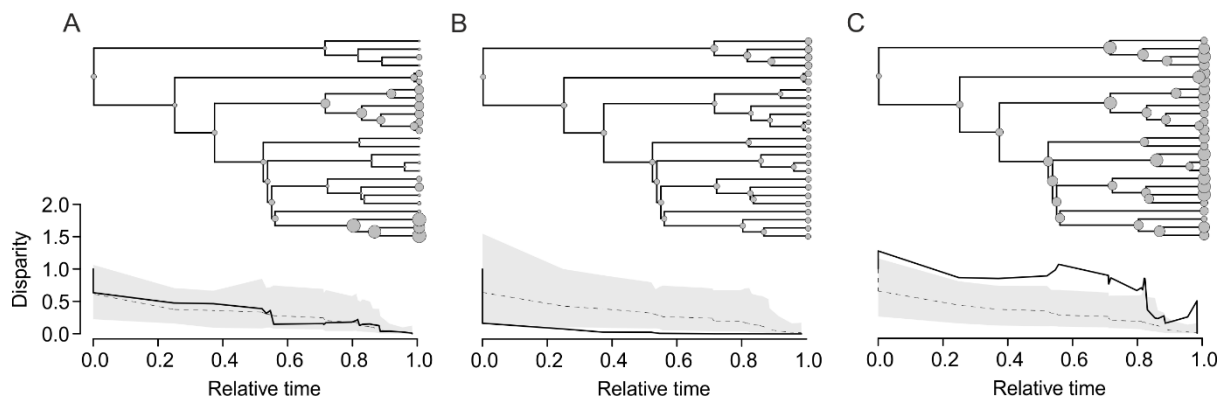
1841
1842 **Fig. 13.** Tracing character ('trait') evolution along a phylogeny. Either categorical (pie charts) or continuous (grey-
1843 step gradient) character states can be estimated along the branches of a phylogeny solely based on trait information
1844 of extant species. Uncertainties of character tracing are here exemplified for the categorical states by the
1845 proportions of the pie charts, but uncertainty can also be estimated for continuous traits.

1846
1847 The second interest concerns modeling tempo (i.e., trait divergence per time) and mode (e.g.,
1848 adaptive, non-adaptive, divergence, convergence) of mainly continuous character evolution. The
1849 underlying assumption is that these two parameters might not be constant over time but are potentially
1850 driven by extrinsic factors (*sensu* Schluter, 2000; Coyne and Orr, 2004). This assumption offers the
1851 possibility to test the effects of geological, climatic, and/or environmental changes in ancient lakes
1852 inferred from sediment cores on species traits over time using specific models. Thereby, a major goal of
1853 interdisciplinary deep-drilling projects can be addressed – understanding the consequences of
1854 environmental change for the biotic evolution of ancient lake taxa as well as the capacity of an ancient
1855 lake to buffer such perturbations.

1856 Three testable models are particularly useful for ancient lake studies (Fig. 14). The simplest model
1857 ('Brownian motion'; Edwards and Cavalli-Sforza, 1964), assumes a constant trait mean of species over
1858 time, which would indicate that extrinsic factors have no effect on character evolution (Fig. 14A).
1859 Because of its neutral characteristic, it is often used as null model to compare to empirical data.

1860 The 'early burst model' assumes that character variance decreases exponentially with time (Blomberg
1861 et al., 2003; Harmon et al., 2010; Fig. 14B). It can, for example, be used to test whether the formation of
1862 an ancient lake triggered the onset of adaptive radiations (*sensu* Schluter, 2000). The third, the 'Hansen
1863 model', allows species to be influenced by past environmental events but assumes that the trait mean
1864 adapts to an optimum over time (Fig. 14C; Hansen, 1997). This model focusses on the relationship
1865 between environment and the pull towards an optimal character state ('selection'; see Hansen, 2014).

1866 Potential changes over time in tempo and mode of character evolution can be visualized by disparity-
1867 through-time plots (Harmon et al., 2003; see Muschick et al., 2012 for an example from Lake
1868 Tanganyika), displaying mean pairwise differences of traits among all species of one clade that is present
1869 at a particular moment in time (Fig. 14).



1870

1871

1872 **Fig. 14.** Simulation of three modes (A–C) of trait evolution (grey circles) showing their signatures in the respective
 1873 disparity-through-time plots (trait divergence $\sigma^2 = 0.3$). Observed traits at the tips of the phylogeny are used to
 1874 estimate the 95% confidence interval over time (grey area). Deviation of the observed disparity (solid black line)
 1875 from this expectancy may indicate periods of environmental influence on evolution or changes in selection strength.
 1876 A) Brownian motion, which is a model of phylogeny-wide stasis in trait mean and the null expectancy in
 1877 evolutionary studies. B) Early burst of trait evolution with decreasing trait variance over time. C) Hansen model of
 1878 trait evolution with trait mean approaching an optimum.

1879

1880 However, tracing character evolution is subject to several pitfalls and challenges. Phylogeny-based
 1881 analyses of character evolution through time can be biased as morphological exuberant species may be
 1882 more prone to extinction (Huang et al., 2015). The effects of missing extinct species on phylogeny-based
 1883 character analysis has not been studied comprehensively yet (but see Albert et al., 2009; Slater et al.,
 1884 2012). However, fossil species may provide complementary insights into character evolution (e.g., Van
 1885 Bocxlaer and Hunt, 2013; Benson et al., 2014) and could be used to fine-tune or even verify models of
 1886 character evolution (Betancur-R et al., 2015). Another challenge is that information on species traits and
 1887 detailed information on morphology, ecological attributes, and distribution of the taxa of concern is often
 1888 sparse. Quality issues include sampling completeness, taxonomic coverage, the presence of cryptic
 1889 species, and DNA marker choice, all affecting the quality of phylogenetic trees (node support,
 1890 resolution) and subsequent state reconstructions (ambiguous states).

1891 Examples of tracing character evolution in ancient lakes include studies on the effect of lake origin on
 1892 changes in morphological traits. Gonzalez-Voyer and Kolm (2011), for example, showed two periods of
 1893 increased morphological disparity in Lake Tanganyika, the older was related to the initial colonization

1894 after lake origination but the younger one could not be explained by limnological history. Tracing
1895 character evolution was also used to infer trophic specialization and coloration patterns of endemic
1896 invertebrates of the ancient Malili Lakes of Sulawesi, Indonesia (von Rintelen et al., 2004, 2010) and to
1897 uncover convergent evolution in ancient lakes (e.g., Meixner et al., 2007; Young et al., 2009). The latter
1898 phenomenon has been puzzling evolutionary biologists for years (e.g., Mahler et al., 2013) and might be
1899 of eminent importance for the interpretation of fossil records from sediment cores. Other applications
1900 include reconstructing the colonization history of ancient lakes over time (e.g., Van Bocxlaer et al., 2015;
1901 Daniels et al., 2015), habitat or niche occupation of ancient lake species such as the colonization of rocky
1902 habitats in Lake Tanganyika by cichlids (Koblmüller et al., 2004), or the bathymetric range evolution of
1903 limpet gastropods in lakes Baikal (Stelbrink et al., 2015) and Ohrid (Albrecht et al., 2006).

1904 From a methodological point of view, there are three trends in tracing character evolution that might
1905 be of importance for future deep drilling projects. First, phylogenetic uncertainties (Sorenson et al., 2014;
1906 Shi and Rabosky, 2015) and character variance caused by measurement errors or intraspecific variation
1907 (Revell, 2012; Clavel et al., 2015) need to be considered. Second, shifts in the tempo (i.e., different rates
1908 of trait divergence; Eastman et al., 2011; Thomas and Freckleton, 2012) or mode (e.g., from neutral
1909 divergence to adaptation; Clavel et al., 2015) of character evolution over time are to be identified, which
1910 may help mitigating erroneous ancestral state reconstruction (King and Lee, 2015). Third, characters
1911 may not evolve independently of each other but co-vary or even constrain one another. Therefore,
1912 multivariate evolutionary models that simultaneously use a set of characters with several states each are
1913 being designed (Freckleton, 2012; Mahler et al., 2013; Adams and Collyer, 2015; Clavel et al., 2015).

1914 Whereas most earlier attempts did not integrate geological and evolutionary data from ancient lakes
1915 at once, such integration is becoming increasingly important in analyses of character changes. Integrated
1916 approaches allow to test whether geological or environmental changes previously inferred from deep
1917 drilling campaigns had an effect on patterns of character change in extant species (e.g., Danley et al.,
1918 2012; Van Bocxlaer and Hunt, 2013; Lyons et al., 2015; Stelbrink et al., 2015). Potential drivers of
1919 previously inferred changes in species traits can also be studied with data that are subsequently generated
1920 from deep drilling campaigns (e.g., von Rintelen et al., 2010 for Lake Towuti and Trajanovski et al.,
1921 2010 for Lake Ohrid).

1922 For future deep drilling campaigns, we expect to see more explicit attempts to link geological and
1923 biological histories. As more phylogenetic data become available, future comparative analyses across
1924 taxa will help shed light on general evolutionary processes in ancient lakes that affect entire communities
1925 (e.g., Salzburger et al., 2014). If common temporal signals in phylogenies can be detected (e.g.,
1926 simultaneous character or rate changes across taxa; O'Meara et al., 2006), a more straightforward link
1927 between environmental changes and evolutionary patterns may be established. These signals could also
1928 help to understand the biological consequences of environmental change in ancient lakes, even when
1929 fossil information is absent.

1930 **4. Conclusions**

- 1931 1) Over the past years, scientific drilling projects in ancient lakes became increasingly interdisciplinary
1932 and have intensified the use of secondary data, i.e., data obtained independently of the drilling
1933 operation. Comprehensive interdisciplinary projects enable a more holistic view on scientific
1934 problems and provide excellent opportunities for hypothesis-driven research.
- 1935 2) One of the most challenging tasks for answering novel research questions in deep drilling projects is
1936 to link diverse datasets with different resolutions, different data qualities, and potentially different
1937 age uncertainties to solve complex problems.
- 1938 3) Careful consideration of drill sites and drilling strategies are a prerequisite to optimize the chances
1939 that the goals of a deep drilling project can be reached. Meticulous preparation, including the
1940 collection of site-specific information from pilot studies may serve to construct a strategy for
1941 collecting primary and secondary data that can then be evaluated against scientific objectives,
1942 budget, logistic requirements, and the available time.
- 1943 4) Accurate subsampling and data interpretation requires a full tracking record of the core. Moreover,
1944 sediment subsampling should be coordinated and samples should, whenever possible, be taken from
1945 the same sample depth to aid data integration.
- 1946 5) Data analyses should be hierarchically structured. As studies from long records are very laborious
1947 and time-consuming, the first target should be to produce low/medium resolution data. Only then,

1948 high-resolution analyses are to be conducted.

1949 6) Though many new methods and analyses (e.g., analysis of isotopes in organic materials, high
1950 throughput image analyses and high throughput DNA sequencing, low-concentration biomarker
1951 analyses, studies of ancient amino acids and ancient DNA, molecular dating) are of high potential
1952 for integrating diverse datasets, their weaknesses in terms of resolution, quality of data, and
1953 practicability have to be considered.

1954 7) Whereas the physical linkage of primary information obtained from sediment cores is, in most cases,
1955 straightforward due to the chronological constraints on the data, integrating secondary data and/or
1956 interpretations into a consistent representation of the natural history of ancient lakes remains
1957 challenging.

1958 8) Moreover, in most previous deep drilling projects, geological and biological data were linked
1959 empirically. However, recent statistical developments enable a better mathematical integration of
1960 diverse types of datasets and the testing of hypotheses based on specific null models.

1961 9) For future deep drilling campaigns we expect to see more explicit attempts to statistically link
1962 geological and biological histories aided by methodological advances in data generation (e.g.,
1963 automated methods of fossil retrieval, visualization and identification) and data analyses (e.g., a
1964 better integration of uncertainties in age-depth calculations).

1965 10) Interdisciplinary projects should integrate earth and life scientists, statisticians, and modelers in the
1966 planning phase, to establish clear communication strategies, to align interests, and to discuss data
1967 requirements.

1968 11) Finally, the interdisciplinary character of modern deep drilling projects not only requires a
1969 constant adaptation to methodological innovations, but also targeted scientific training
1970 components. Therefore, thematic workshops, network training events and/or field schools,
1971 particularly for early stage researchers, should be an integrative part of interdisciplinary deep
1972 drilling campaigns.

1973 **Glossary**

- 1974 *Accommodation space*: Available space for accumulation of sediments.
- 1975 *Adaptive radiation*: Rapid diversification of species accompanied by adaptation into various niches. The
- 1976 term is used both to describe an evolutionary process as well as the result of this process.
- 1977 *Age-depth model*: Synthetic model that explains the relationship between sediment depth and sediment
- 1978 age in depositional environments.
- 1979 *Allopatric speciation* (= *geographical speciation*): Speciation due to the evolution of (geographical)
- 1980 reproductive barriers in populations that prevent or interfere with gene flow.
- 1981 *Amino acid racemization*: Spontaneous reaction describing the interconversion between the chiral forms
- 1982 of an amino acid.
- 1983 *Aminostratigraphy*: Relative dating framework based on the extent of amino acid racemization in
- 1984 subfossil biominerals.
- 1985 *Anagenesis*: Directional evolutionary change from an ancestor species to a descendant species without
- 1986 lineage splitting (also see chronospecies).
- 1987 *Ancestral polymorphism*: The amount of heterogeneity that is present in an ancestral population prior to
- 1988 the separation of the descending species.
- 1989 *Ancestral state reconstruction*: Reconstruction of ancestral phenotypic or genetic states of species along
- 1990 a phylogenetic tree based on information of extant taxa.
- 1991 *Ancient lake*: A lake that has continuously existed for > 100 ky or even > 1 My. The meaning of the term
- 1992 is not universally accepted. Some authors use this term synonymously with ‘long-lived lake’. Others
- 1993 use the term ancient lake only for extant long-lived lakes.
- 1994 *Ancient lake species flock*: Species rich, monophyletic group of endemic taxa that typically evolved
- 1995 within the lake (i.e., intralacustrine).
- 1996 *Bioindicator*: Extant species that are used to infer the present ecological conditions of an ecosystem.
- 1997 *Biomarker, sedimentary*: Source-specific organic molecules (‘molecular fossils’).
- 1998 *Biovolume accumulation rate*: Diatom concentration divided by sediment accumulation rate and
- 1999 corrected for size differences between taxa.

2000 *Borehole logging*: Process of measuring physical, chemical, and structural properties of penetrated
2001 geological formations using logging tools that are lowered into a borehole on a wireline cable.

2002 *Bottleneck, genetic*: Sudden decrease in population size, which potentially reduces the genetic variation
2003 within a population.

2004 *Coalescence analyses*: Population genetic analyses that relate patterns of genetic diversity in an extant
2005 population to its demographic history.

2006 *Chronospecies*: Arbitrary divisions of a single evolutionary lineage, defined on the basis of
2007 morphological change within the lineage (also see anagenesis).

2008 *Cladogenesis*: Evolutionary branching of an ancestor species into two or more descendant species.

2009 *Composite core*: Layer-to-layer correlation of core segments from multiple boreholes drilled at the same
2010 drill site, i.e., best-case scenario of a continuous, undisturbed sediment profile.

2011 *Convergence*: Similarities that have arisen independently in two or more organisms that do not share a
2012 common ancestry.

2013 *Cyclostratigraphy*: Study of stratigraphic records of astronomically forced climate cycles.

2014 *Depocenter*: Location of the thickest deposit in a sedimentary basin.

2015 *Divergence time*: Time since separation of descendent taxa from a most recent common ancestor.

2016 *Endemism*: Characteristic of a taxon that is restricted to a geographic location (such as an ancient lake).

2017 *Evolution, biological*: Change in heritable traits of populations from generation to generations.

2018 *Evolution, biotic*: Gradual change in the structure, composition, or dynamics of biological objects or
2019 systems.

2020 *Evolution, geological*: Gradual change in the structure, composition, or dynamics of geological objects
2021 or systems.

2022 *Global benthic isotope stack*: Stack of 57 benthic marine $\delta^{18}\text{O}$ records reflecting global ice volume and
2023 deep ocean temperature for the past 5.3 My, often used as stratigraphic reference record.

2024 *Hiatus*: Discontinuity ('break') in the stratigraphic succession.

2025 *Incomplete lineage sorting*: Phenomenon that not all genetic lineages are segregated at the time of
2026 species splitting.

2027 *Interdisciplinarity*: Integration of two or more scientific disciplines.

2028 *Intralacustrine*: Within a lake.

2029 *Lineages-through-time plot*: Plot showing the accumulation of lineages through time in a time-calibrated
2030 phylogeny.

2031 *Lake proper*: Lake body excluding peripheral water bodies or effluents/affluents.

2032 *Long-lived lake*: A lake that has continuously existed for > 100 ky or even > 1 My. Some authors use this
2033 term synonymously with the term ‘ancient lake’.

2034 *Macrofossils*: Remains of organisms from the remote past large enough to be visible without a
2035 microscope.

2036 *Magnetic susceptibility*: A measure of the degree of magnetization of a material after the application of a
2037 magnetic field.

2038 *Microbiome*: It refers to the entire microbial population within a specific environmental niche.

2039 *Microfossils*: Microscopically small remains of organisms from the remote past.

2040 *Molecular clock*: A concept that correlates number of nucleotide or amino acid substitutions (i.e.,
2041 mutations fixed in the genome) to time.

2042 *Multidisciplinarity*: Concurrent combination of two or more scientific disciplines.

2043 *Phylogenetic tree*: Graphical representation of evolutionary (genealogical) relationships of several
2044 species or other units, which are assumed to have a common ancestor.

2045 *Pollen rain*: The cloud of airborne pollen produced by plants.

2046 *Proxy*: Measured variable used to model or generate the value of a variable that is typically more
2047 difficult to obtain.

2048 *Radiation, evolutionary*: Event of rapid cladogeneses.

2049 *Relaxed clock*: A dating approach that relaxes the assumption of a single substitution rate within a
2050 phylogeny and allows rates to vary across the branches.

2051 *Resilience, ecosystem*: Ability of an ecosystem to resist disturbances.

2052 *Speciation*: Evolutionary process leading to new species.

2053 *Species flock*: In ancient lakes, monophyletic group of endemic species that evolved intralacustrine.

2054 *Stable isotope*: An isotope of an element that does not tend to decay over time.

2055 *Substitution rate*: Here used in terms of number of fixed mutations per site and time unit.

2056 *Subsurface biosphere*: Term used to designate the active microbial life in lacustrine sediments in analogy
2057 to deep biosphere that refers to the marine environment.

2058 *Tephra*: Pyroclastic material ejected from a volcano including fragmented rocks and smaller particles.

2059 *Trait*: An inherited morphological, molecular, or ecological characteristic of a species.

2060 *Trait-specific clock*: A single molecular clock rate of a specific gene that can be assigned to a range of
2061 taxa that share similar biological and life history characteristics that are supposedly affecting rate
2062 heterogeneity.

2063 *Transfer functions, paleoecological*: Models that correlate modern species assemblages and their
2064 environmental characteristics to fossil assemblages for reconstructing past environmental conditions.

2065 *Watershed*: Catchment area of a drainage basin.

2066 **Acknowledgements**

2067 This work was supported by German Research Foundation (DFG) grants WA 2109/11, WI 1902/7,
2068 WI1902/13, and AL 1076/6, 9 to B. Wagner, T. Wilke, and C. Albrecht, respectively, by the European
2069 Commission, Marie Skłodowska-Curie Action Innovative Training Network 642973 ‘Pontocaspian Rise
2070 and Demise (PRIDE)’ to F. Wesselingh and T. Wilke, by an Alexander von Humboldt Fellowship to B.
2071 Van Bocxlaer, by Swiss National Science Foundation (SNSF) grants 20FI21_153054/1 and
2072 200021_153053/1 awarded to H. Vogel and D. Ariztegui as well as SNSF grants 200020-119931/2,
2073 200021-132529 and 200020-149221/1 to D. Ariztegui. Many of the results presented here have been
2074 possible through the financial support of various ICDP-sponsored projects. We also thank the editor and
2075 two anonymous referees for valuable and constructive comments that significantly improved the paper.

2076

2077 **References**

- 2078 Abarca, N., Jahn, R., Zimmermann, J., Enke, N., 2014. Does the cosmopolitan diatom *Gomphonema parvulum*
2079 (Kützing) Kützing have a biogeography? PLoS ONE 9, e86885.
2080 <http://dx.doi.org/10.1371/journal.pone.0086885>.
- 2081 Adams, D. C., Collyer, M. L., 2015. Permutation tests for phylogenetic comparative analyses of high-dimensional
2082 shape data: What you shuffle matters. *Evolution* 69, 823-829. <http://dx.doi.org/10.1111/evo.12596>.
- 2083 Adams, D. C., 2008. Phylogenetic meta-analysis. *Evolution* 62, 567-572. <http://dx.doi.org/10.1111/j.1558->
2084 [5646.2007.00314.x](http://dx.doi.org/10.1111/j.1558-5646.2007.00314.x).
- 2085 Aichner, B., Feakins, S. J., Lee, J. E., Herzsuh, U., Liu, X., 2015. High-resolution leaf wax carbon and hydrogen
2086 isotopic record of the late Holocene paleoclimate in arid Central Asia. *Clim. Past* 11, 619-633.
2087 <http://dx.doi.org/10.5194/cp-11-619-2015>.
- 2088 Albert, J. S., Johnson, D. M., Knouft, J.H., 2009. Fossils provide better estimates of ancestral body size than do
2089 extant taxa in fishes. *Acta Zool.* 90, 357-384. <http://dx.doi.org/10.1111/j.1463-6395.2008.00364.x>.
- 2090 Albert, P. G., Hardiman, M., Keller, J., Tomlinson, E. L., Smith, V. C., Bourne, A. J., Wulf, S., Zanchetta, G.,
2091 Sulpizio, R., Müller, U. C., Pross, J., Ottolini, L., Matthews, I. P., Blockley, S. P. E., Menzies, M. A., 2015.
2092 Revisiting the Y-3 tephrostratigraphic marker: a new diagnostic glass geochemistry, age estimate, and details on
2093 its climatostratigraphical context. *Quat. Sci. Rev.* 118, 105-121.
2094 <http://dx.doi.org/10.1016/j.quascirev.2014.04.002>.
- 2095 Albrecht, C., Trajanovski, S., Kuhn, K., Streit, B., Wilke, T., 2006. Rapid evolution of an ancient lake species
2096 flock: freshwater limpets (Gastropoda: Ancyliidae) in the Balkan Lake Ohrid. *Organisms Diversity & Evolution*
2097 6, 294-307.
- 2098 Albrecht, C., Wilke, T., 2008. Ancient Lake Ohrid: biodiversity and evolution. *Hydrobiologia* 615: 103-140.
- 2099 Allentoft, M. E., Collins, M., Harker, D., Haile, J., Oskam, C. L., Hale, M. L., Campos, P. F., Samaniego, J. A.,
2100 Gilbert, M. T., Willerslev, E., Zhang, G., Scofield, R. P., Holdaway, R. N., Bunce, M., 2012. The half-life of
2101 DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc. R. Soc. Lond. B* 279, 4724-4733.
2102 <http://dx.doi.org/10.1098/rspb.2012.1745>.
- 2103 Anderson-Carpenter, L. L., McLachlan, J. S., Jackson, S. T., Kuch, M., Lumibao, C. Y., Poinar, H. N., 2011.
2104 Ancient DNA from lake sediments: bridging the gap between paleoecology and genetics. *BMC Evol. Biol.* 11,
2105 30.

- 2106 Ariztegui, D., Thomas, C., Vuillemin, A., 2015. Present and future of subsurface biosphere studies in lacustrine
2107 sediments through scientific drilling. *Int. J. Earth Sci. (Geol. Rundsch.)* 104, 1655-1665.
2108 <http://dx.doi.org/10.1007/s00531-015-1148-4>.
- 2109 Avise, J. C., 2000. *Phylogeography. The history and formation of species*. Harvard University Press, Cambridge,
2110 Massachusetts, USA & London, U.K.
- 2111 Ayala, F. J., 1997. Vagaries of the molecular clock. *P. Natl. Acad. Sci. USA* 94, 7776-7783.
- 2112 Barnosky, A. D., 2001. Distinguishing the effects of the Red Queen and Court Jester on Miocene mammal
2113 evolution in the northern Rocky Mountains. *J. Vertebr. Paleontol.* 21, 172-185.
- 2114 Battarbee, R. W., 1986. Diatom analysis. In: Berglund, B. E. (Ed.), *Handbook of Holocene palaeoecology and*
2115 *palaeohydrology*. John Wiley & Sons, Chichester, pp. 527-570.
- 2116 Baumgarten, H., Wonik, T., 2014. Cyclostratigraphic studies of sediments from Lake Van (Turkey) based on their
2117 uranium contents obtained from downhole logging and paleoclimatic implications. *Int. J. Earth Sci.*, 1-16.
- 2118 Baumgarten, H., Wonik, T., Tanner, D. C., Francke, A., Wagner, B., Zanchetta, G., Sulpizio, R., Giaccio, B.,
2119 Nomade, S., 2015. Age-depth model of the past 630 kyr for Lake Ohrid (Macedonia/Albania) based on
2120 cyclostratigraphic analysis of downhole gamma ray data. *Biogeosci.* 12, 7453-7465.
2121 <http://dx.doi.org/10.5194/bg-12-7453-2015>.
- 2122 Bell, M. A., Travis, M. P., Blouw D. M., 2006. Inferring natural selection in a fossil threespine stickleback.
2123 *Paleobiology* 32, 446-464.
- 2124 Bell, M. A., Sadagursky, M. S., Baumgartner J. V., 1987. Utility of lacustrine deposits for the study of variation
2125 within fossil samples. *Palaios* 2, 455-466.
- 2126 Benavides, E., 2005. The *Telmatobius* species complex in Lake Titicaca: applying phylogeographic and coalescent
2127 approaches to evolutionary studies of highly polymorphic Andean frogs. *Monogr. Herpetol.* 7, 167-185.
- 2128 Benson, R. B. J., Campione, N. E., Carrano, M. T., Mannion, P. D., Sullivan, C., Upchurch, P., Evans, D. C., 2014.
2129 Rates of Dinosaur Body Mass Evolution Indicate 170 Million Years of Sustained Ecological Innovation on the
2130 Avian Stem Lineage. *PLoS Biol.* 12, e1001853. <http://dx.doi.org/10.1371/journal.pbio.1001853>.
- 2131 Berglund, B. E., Ralska-Jasiewiczowa, M., 1986. Pollen analysis and pollen diagrams. In: Berglund, B. E. (Ed.),
2132 *Handbook of Holocene Palaeoecology and Palaeohydrology*. John Wiley & Sons, Chichester, pp. 455-496.
- 2133 Bergner, A. G. N., Trauth, M. H., 2004. Comparison of the hydrological and hydrochemical evolution of Lake
2134 Naivasha (Kenya) during three highstands between 175 and 60 kyr BP. *Palaeogeogr. Palaeoclimatol.*
2135 *Palaeoecol.* 215, 17-36. <http://dx.doi.org/10.1016/j.palaeo.2004.07.033>.

- 2136 Berke, M. A., Johnson, T. C., Werne, J. P., Grice, K., Schouten, S., Sinninghe Damsté, J. S., 2012. Molecular
2137 records of climate variability and vegetation response since the Late Pleistocene in the Lake Victoria basin, East
2138 Africa. *Quat. Sci. Rev.* 55, 59-74.
- 2139 Bertini, A., 2010. Pliocene to Pleistocene palynoflora and vegetation in Italy: State of the art. *Quat. Int.* 225, 5-24.
- 2140 Betancur-R, R., Ortí, G., Pyron, R. A., 2015. Fossil-based comparative analyses reveal ancient marine ancestry
2141 erased by extinction in ray-finned fishes. *Ecol. Lett.* 18, 441-450. <http://dx.doi.org/10.1111/ele.12423>.
- 2142 Birks, H. H., Birks, H. J. B., 2000. Future uses of pollen analysis must include plant macrofossils. *J. Biogeogr.* 27,
2143 31-35.
- 2144 Birks, H. H., Birks, H. J. B., 2006. Multi-proxy studies in palaeolimnology. *Veg. Hist. Archaeobot.* 15, 235-251.
- 2145 Birks, H. J. B., Birks, H. H. 2016. How have studies of ancient DNA from sediments contributed to the
2146 reconstruction of Quaternary floras? *New Phytol.* 209, 499-506. <http://dx.doi.org/10.1111/nph.13657>.
- 2147 Birks, H. J. B., Lotter, A. F., Juggins, S., Smol, J. P. (Eds.), 2012. Tracking environmental change using lake
2148 sediments: data handling and numerical techniques, Vol. 5. Springer Netherlands.
- 2149 Bissett, A., Gibson, J. A. E., Jarman, S. N., Swadling, K. M., Cromer, L., 2005. Isolation, amplification, and
2150 identification of ancient copepod DNA from lake sediments. *Limnol. Oceanogr. Methods* 3, 533-542.
- 2151 Blaauw, M., 2010. Methods and code for 'classical' age-modelling of radiocarbon sequences. *Quat. Geochronol.* 5,
2152 512-518. <http://dx.doi.org/10.1016/j.quageo.2010.01.002>.
- 2153 Blaauw, M., Christen, J. A., 2011. Flexible paleoclimate age-depth models using an autoregressive gamma process.
2154 *Bayesian Anal.* 6, 457-474. <http://dx.doi.org/10.1214/ba/1339616472>.
- 2155 Blaauw, M., Heegaard, E., 2012. Estimation of age-depth relationships. In: Birks, H. J. B., Juggins, S., Lotter, A.,
2156 Smol, J. P. (Eds.), *Tracking environmental change using lake sediments*. Springer, Dordrecht, pp. 379-413.
2157 http://dx.doi.org/10.1007/978-94-007-2745-8_12.
- 2158 Bloemsa, M. R., Zabel, M., Stuut, J. B. W., Tjallingii, R., Collins, J. A., Weltje, G. J., 2012. Modelling the joint
2159 variability of grain size and chemical composition in sediments. *Sedimentary Geology* 280, 135-148.
- 2160 Blomberg, S. P., Garland, T., Ives, A. R., 2003. Testing for Phylogenetic Signal in Comparative Data: Behavioral
2161 Traits Are More Labile. *Evolution* 57, 717-745. <http://dx.doi.org/10.1111/j.0014-3820.2003.tb00285.x>.
- 2162 Boessenkool, S., McGlynn, G., Epp, L. S., Taylor, D., Pimentel, M., Gizaw, A., Nemomissa, S., Brochmann, C.,
2163 Popp, M., 2014. Use of ancient sedimentary DNA as a novel conservation tool for high-altitude tropical
2164 biodiversity. *Conserv. Biol.* 28, 446-455. <http://dx.doi.org/10.1111/cobi.12195>.
- 2165 Bond, G., Heinrich, H., Broecker, W., Labeyrie, L., McManus, J., Andrews, J., Huon, S., Jantschik, R., Clasen, S.,

2166 Simet, C., Tedesco, K., Klas, M., Bonani, G., Ivy, S., 1992. Evidence for massive discharges of icebergs into the
2167 North Atlantic Ocean during the last glacial period. *Nature* 360, 245-249.

2168 Bond, G. C., Lotti, R., 1995. Iceberg discharges into the North Atlantic on millennial time scales during the last
2169 glaciation. *Science* 267, 1005-1010.

2170 Bowen, D. Q., Hughes, S. A., Sykes, G. A., Miller, G. H., 1989. Land-sea correlations in the Pleistocene based on
2171 isoleucine epimerisation in non-marine molluscs. *Nature* 340, 49-51.

2172 Bowen G. J., Wilkinson, B., 2002. Spatial distribution of $\delta^{18}\text{O}$ in meteoric precipitation. *Geology* 30, 315-318.

2173 Bradbury, P. J., 1999. Continental diatoms as indicators of long-term environmental change. In: Stoermer, E. F.,
2174 Smoll, J. P. (Eds.), *The Diatoms: Applications for the Environmental and Earth Sciences*. Cambridge University
2175 Press, Cambridge, pp. 169-182.

2176 Bradley, R. S., 2014. *Paleoclimatology: Reconstructing Climate of the Quaternary*. Elsevier Academic Press, San
2177 Diego.

2178 Brocks, J. J., Grosjean, E., Logan, G. A., 2008. Assessing biomarker syngeneity using branched alkanes with
2179 quaternary carbon (BAQCs) and other plastic contaminants. *Geochim. Cosmochim. Acta* 72, 871-888.

2180 Broecker, W. S., 1994. Massive iceberg discharges as triggers for global climate change. *Nature* 372, 421-424.

2181 Brooks, A. S., Hare, P. E., Kokis, J. E., Miller, G. H., Ernst, R. D., Wendorf, F., 1990. Dating Pleistocene
2182 archaeological sites by protein diagenesis in ostrich eggshell. *Science* 248, 60-64.

2183 Brooks, J. L., 1950. Speciation in ancient lakes. *Quat. Rev. Biol.* 25, 30-60, 131-176.

2184 Brooks, J., Shaw, G., 1978. Sporopollenin: A review of its chemistry, palaeochemistry and geochemistry. *Grana* 17,
2185 91-97.

2186 Brooks, P. W., Eglinton, G., Gaskell, S. J., McHugh, D. J., Maxwell, J. R., Philp, R. P., 1976. Lipids of recent
2187 sediments. Part 1: straight chain hydrocarbons and carboxylic acids of some temperate lacustrine and
2188 subtropical lagoonal/tidal fiat sediments. *Chem. Geol.* 18, 21-38.

2189 Brown, K. J., Rüber, L., Bills, R., Day, J. J., 2010. Mastaceblid eels support Lake Tanganyika as an evolutionary
2190 hotspot of diversification. *BMC Evol. Biol.* 10, 188. <http://dx.doi.org/10.1186/1471-2148-10-188>.

2191 Buckles, L. K., Weijers, J. W. H., Verschuren, D., Sinnighe Damsté, J. S., 2014. Sources of core and intact
2192 branched tetraether membrane lipids in the lacustrine environment: Anatomy of Lake Challa and its catchment,
2193 equatorial East Africa. *Geochim. Cosmochim. Acta* 140, 106-126.

2194 Buckley, M., Wadsworth, C., 2014. Proteome degradation in ancient bone: Diagenesis and phylogenetic potential.
2195 *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 416, 69-79.

2196 Buckley, M., Walker, A., Ho, S. Y., Yang, Y., Smith, C., Ashton, P., Thomas-Oates, J., Cappellini, E., Koon, H.,
2197 Penkman, K., Elsworth, B., Ashford, D., Solazzo, C., Andrews, P., Strahler, J., Shapiro, B., Ostrom, P., Gandhi,
2198 H., Miller, W., Raney, B., Zylber, M. I., Gilbert, M. T., Prigodich, R. V., Ryan, M., Rijdsdijk, K. F., Janoo, A.,
2199 Collins, M. J., 2008. Comment on "Protein Sequences from Mastodon and Tyrannosaurus rex Revealed by
2200 Mass Spectrometry". *Science* 319, 33c-33c.

2201 Bundy, J. G., Davey, M. P., Viant, M. R., 2009. Environmental metabolomics: a critical review and future
2202 perspectives. *Metabolomics* 5, 3-21.

2203 Bush, A. W., Powell, M. G., Arnold, W. S., Bert, T. M., Daley, G. M., 2002. Time-averaging, evolution, and
2204 morphologic variation. *Paleobiology* 28, 9-25.

2205 Camel, V., 2000. Microwave-assisted solvent extraction of environmental samples. *Trends Anal. Chem.* 19, 229-
2206 248.

2207 Cantalapiedra, J. L., FitzJohn, R. G., Kuhn, T. S., Fernández, M. H., DeMiguel, D., Azanza, B., Morales, J.,
2208 Mooers, A. Ø., 2014. Dietary innovations spurred the diversification of ruminants during the Cenozoic. *Proc.*
2209 *R. Soc. Lond. B* 281, 20132746. <http://dx.doi.org/10.1098/rspb.2013.2746>.

2210 Cappellini, E., Jensen, L. J., Szklarczyk, D., Ginolhac, A., Da Fonseca, R. A. R., Stafford, T., Holen, S. R., Collins,
2211 M. J., Orlando, L., Willerslev, E., Gilbert, M. T. P., Olsen, J. V., 2011. Proteomic analysis of a Pleistocene
2212 mammoth femur reveals more than one hundred ancient bone proteins. *J. Proteome Res.* 11, 917-926.

2213 Castañeda, I. S., Schouten, S., 2011. A review of molecular organic proxies for examining modern and ancient
2214 lacustrine environments. *Quat. Sci. Rev.* 30, 2851-2891.

2215 Cavalli-Sforza, L. L., Edwards, A. W. F., 2005 Reconstruction of evolutionary trees. In: Heywood V. H., McNeill
2216 J. (Eds.) *Phenetic and Phylogenetic Classification*. London: Systematics Association Publications No. 6,
2217 London, pp. 67-76.

2218 Clark, I., Fritz, I., 1997. *Environmental Isotopes in Hydrogeology*. Lewis, Boca Raton.

2219 Clavel, J., Escarguel, G., Merceron, G., 2015. mvMORPH: an R package for fitting multivariate evolutionary
2220 models to morphometric data. *Methods Ecol. Evol.* 6, 1311-1319. <http://dx.doi.org/10.1111/2041-210X.12420>.

2221 Coates, G. R., Xiao, L., Prammer, M. G., 1999. *NMR Logging: Principles and Applications*. Haliburton Energy
2222 Services, Houston.

2223 Cocquyt, C., 1998. Diatoms from the Northern Basin of Lake Tanganyika. *Biblioth. Diatomol.* 39, 1-276.

2224 Cohen, A. S., 2003. *Paleolimnology: the history and evolution of lake systems*. Oxford University Press, New
2225 York.

- 2226 Cohen, A. S., 2012. Scientific drilling and biological evolution in ancient lakes: lessons learned and
 2227 recommendations for the future. *Hydrobiologia* 682, 3-25. <http://dx.doi.org/10.1007/s10750-010-0546-7>.
- 2228 Cohen, A. S., Stone, J. R., Beuning, K. R. M., Park, L. E., Reinthal, P. N., Dettman, D., Scholz, C. A., Johnson, T.
 2229 C., King, J. W., Talbot, M. R., Brown, E. T., Ivory, S. J., 2007. Ecological consequences of early Late
 2230 Pleistocene megadroughts in tropical Africa. *Proc. Natl. Acad. Sci. USA* 104, 16422-16427.
 2231 <http://dx.doi.org/10.1073/pnas.0703873104>.
- 2232 Collins, M., Demarchi, B., 2014. Amino Acid Racemization, Paleoclimate. In: Rink, W. J., Thompson, J. (Eds.),
 2233 Encyclopedia of scientific dating methods. Springer Netherlands, pp. 1-3.
- 2234 Condamine, F. L., Rolland, J., Morlon, H., 2013. Macroevolutionary perspectives to environmental change. *Ecol.*
 2235 *Lett.* 16, 72-85. <http://dx.doi.org/10.1111/ele.12062>.
- 2236 Costa, K. M., Russell, J. M., Vogel, H., Bijaksana, S., 2015. Hydrological connectivity and mixing of Lake Towuti,
 2237 Indonesia in response to paleoclimatic changes over the last 60,000 years. *Palaeogeogr. Palaeoclimatol.*
 2238 *Palaeoecol.* 417, 467-475.
- 2239 Coyne, J. A., Orr, H. A., 2004. Speciation. Sinauer, Sunderland, Mass., USA.
- 2240 Cranwell, P. A., 1973. Chain-length distribution of *n*-alkanes from lake sediments in relation to post-glacial
 2241 environmental change. *Freshwat. Biol.* 3, 259-265.
- 2242 Craig, H., 1961. Isotopic variations in meteoric waters. *Science* 133, 1833-1834.
- 2243 Craig, H., 1965. The measurement of oxygen isotope palaeotemperatures. In: Tongiorgi, E. (Ed.), *Stable Isotopes in*
 2244 *Oceanographic Studies and Palaeotemperatures*. Pisa, Consiglio Nazionale delle Ricerche Laboratorio di
 2245 *Geologia Nucleare*, pp. 161-182.
- 2246 Cunningham, L., Vogel, H., Nowaczyk, N., Wennrich, V., Juschus, O., Persson, P., Rosén, P., 2013. Climatic
 2247 variability during the last interglacial inferred from geochemical proxies in the Lake El'gygytgyn sediment
 2248 record. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 385, 408-414.
- 2249 Cvetkoska, A., Jovanovska, E., Francke, A., Tofilovska, S., Vogel, H., Levkov, Z., Donders, T., Wagner, B.,
 2250 Wagner-Cremer, F., 2015b. Ecosystem regimes and responses in a coupled ancient lake system from MIS 5b to
 2251 present: The diatom record of lakes Ohrid and Prespa. *Biogeosci. Discuss.* 12, 15051-15086.
- 2252 Cvetkoska, A., Reed, J. M., Levkov, Z., 2012. Diatoms as indicators of environmental change in ancient Lake
 2253 Ohrid during the last glacial–interglacial cycle (ca 140 ka). In: Witkowski, A. (Ed.), *Diatom Monographs*, vol.
 2254 15, ARG Gartner Verlag, Ruggell, Liechtenstein, 220 pp.
- 2255 Daniels, S. R., Phiri, E. E., Klaus, S., Albrecht, C., Cumberlidge, N., 2015. Multi-locus phylogeny of the

- 2256 Afrotropical freshwater crab fauna reveals historical drainage connectivity and transoceanic dispersal since the
2257 Eocene. *Syst. Biol.* 64, 549-567.
- 2258 Danley, P. D., Husemann, M., Ding, B., DiPietro, L. M., Beverly, E. J., Peppe, D. J., 2012. The impact of the
2259 geologic history and paleoclimate on the diversification of East African cichlids. *Int. J. Evol. Biol.* 2012, 1-20.
2260 <http://dx.doi.org/10.1155/2012/574851>.
- 2261 Dansgaard, W., 1964. Stable isotopes in precipitation. *Tellus* 16, 436-468.
- 2262 Dansgaard, W., Jonhson, S. J., Clausen, H. B., Dahl-Jensen, D., Gundestrup, N. S., Hammer, C. U., Hvidberg, C.
2263 S., Steffensen, J. P., Sveinbjornsdottir, A. E., Jouzel, J., Bond, G., 1993. Evidence for general instability of past
2264 climate from a 250-kyr ice-core record. *Nature* 364, 218-220.
- 2265 Davis, M. P., Midford, P. E., Maddison, W., 2013. Exploring power and parameter estimation of the BiSSE method
2266 for analyzing species diversification. *BMC Evol. Biol.* 13, 38. <http://dx.doi.org/10.1186/1471-2148-13-38>.
- 2267 Day, J. J., Cotton, J. A., Barraclough, T. G., 2008. Tempo and mode of diversification of Lake Tanganyika cichlid
2268 fishes. *PLoS ONE* 3, e1730. <http://dx.doi.org/10.1371/journal.pone.0001730>.
- 2269 Dean, J. R., Jones, M. D., Leng, M. J., Noble, S. R., Metcalfe, S. E., Sloane, H. J., Sahy, D., Eastwood, W. J.,
2270 Roberts, C. N., 2015. Eastern Mediterranean hydroclimate over the late glacial and Holocene, reconstructed
2271 from the sediments of Nar lake, central Turkey, using stable isotopes and carbonate mineralogy. *Quat. Sci. Rev.*
2272 124, 162-174.
- 2273 De Decker, P., Last, W. M., 1988. Modern dolomite deposition in continental, saline lakes, western Victoria,
2274 Australia. *Geology* 16, 29-32.
- 2275 De Jonge, C., Hopmans, E. C., Zell, C. I., Kim, J.-H., Schouten, S., Sinninghe Damsté, J. S., 2014. Occurrence and
2276 abundance of 6-methyl branched glycerol dialkyl glycerol tetraethers in soils: Implications for palaeoclimate
2277 reconstruction. *Geochim. Cosmochim. Acta* 141, 97-112.
- 2278 Denis, E. H., Toney, J. L., Tarozo, R., Scott Anderson, R., Roach, L. D., Huang, Y., 2012. Polycyclic aromatic
2279 hydrocarbons (PAHs) in lake sediments record historic fire events: Validation using HPLC-fluorescence
2280 detection. *Org. Geochem.* 45, 7-17.
- 2281 D'Hondt, S., Rutherford, S., Spivack, A. J., 2002. Metabolic activity of subsurface life in deep-sea sediments.
2282 *Science* 295, 2067-2070.
- 2283 Dieleman, J., Van Bocxlaer, B., Manntsche, C., Nyingi, D. W., Adriaens, D., Verschuren, D., 2015. Tracing
2284 functional adaptation in African cichlid fishes through morphometric analysis of fossil teeth: exploring the
2285 methods. *Hydrobiologia* 755, 73-88.

- 2286 Douglas, P. M. J., Pagani, M., Eglinton, T. I., Brenner, M., Hodell, D. A., Curtis, J. H., Ma, K. F., Breckenridge, A.,
2287 2014. Pre-aged plant waxes in tropical lake sediments and their influence on the chronology of molecular
2288 paleoclimate proxy records. *Geochim. Cosmochim. Acta* 141, 346-364.
- 2289 Dumont, H. J., 1998. The Caspian lake: History, biota, structure, and function. *Limnol. Oceanogr.* 43, 44-52.
- 2290 Dunn, K. J., Bergman, D. J., LaTorraca, G. A., 2002. Nuclear Magnetic Resonance – Petrophysical and Logging
2291 Applications. Pergamon, Amsterdam.
- 2292 Dynesius, M., Jansson, R., 2014. Persistence of within-species lineages: a neglected control of speciation rates.
2293 *Evolution* 68, 923-934. <http://dx.doi.org/10.1111/evo.12316>.
- 2294 Eastman, J. M., Alfaro, M. E., Joyce, P., Hipp, A. L., Harmon, L. J., 2011. A Novel Comparative Method for
2295 Identifying Shifts in the Rate of Character Evolution on Trees. *Evolution* 65, 3578-3589.
2296 <http://dx.doi.org/10.1111/j.1558-5646.2011.01401.x>.
- 2297 Edwards, A. W. F., Cavalli-Sforza, L. L., 1964. Reconstruction of evolutionary trees. In: Heywood, V. H., McNeill,
2298 J. (Eds.), Phenetic and phylogenetic classification. Systematics Association pub. no. 6, London, pp. 67–76.
- 2299 Eicher, U., Siegenthaler, U., 1976. Palynological and oxygen isotope investigations on late-Glacial sediment cores
2300 from Swiss lakes. *Boreas* 5, 109-117.
- 2301 Ellis, D. V., Singer, J. M., 2007. Well Logging for Earth Scientists, 2nd ed. Springer, Amsterdam.
- 2302 Faegri, K., Kaland, P. E., Krzywinski, K., 1989. Textbook of pollen Analysis. Wiley, New York.
- 2303 Fazalova, V., Nevado, B., Peretolchina, T., Petunina, J., Sherbakov, D., 2010. When environmental changes do not
2304 cause geographic separation of fauna: differential responses of Baikalian invertebrates. *BMC Evol. Biol.* 10,
2305 320.
- 2306 Filippov, A., Riedel, F., 2009. The late Holocene mollusc fauna of the Aral Sea and its biogeographical and
2307 ecological interpretation. *Limnologica* 39, 67-85.
- 2308 Föller, K., Stelbrink, B., Hauffe, T., Albrecht, C., Wilke, T., 2015. Constant diversification rates of endemic
2309 gastropods in ancient Lake Ohrid: ecosystem resilience likely buffers environmental fluctuations. *Biogeosci.* 12,
2310 7209-7222. <http://dx.doi.org/10.5194/bg-12-7209-2015>.
- 2311 Foote, M., 2000. Origination and extinction components of taxonomic diversity: general problems. *Paleobiology*
2312 26(sp4), 74-102. [http://dx.doi.org/10.1666/0094-8373\(2000\)26\[74:OAECOT\]2.0.CO;2](http://dx.doi.org/10.1666/0094-8373(2000)26[74:OAECOT]2.0.CO;2).
- 2313 Fourtanier, E., Kociolek, J. P., 2011. Catalogue of Diatom Names, California Academy of Sciences.
2314 <http://research.calacademy.org/research/diatoms/names/index.asp>.
- 2315 Francke, A., Wagner, B., Just, J., Leicher, N., Gromig, R., Baumgarten, H., Vogel, H., Lacey, J. H., Sadori, L.,

2316 Wonik, T., Leng, M. J., Zanchetta, G., Sulpizio, R., Giaccio, B., 2016. Sedimentological processes and
2317 environmental variability at Lake Ohrid (Macedonia, Albania) between 640 ka and present day. *Biogeosci.* (in
2318 press).

2319 Frank, U., Nowaczyk, N. R., Negendank, J. F. W., Melles, M., 2002. A paleomagnetic record from Lake Lama,
2320 northern Central Siberia, *Phys. Earth Planet. In.* 133, 3-20.

2321 Freckleton, R. P., 2012. Fast likelihood calculations for comparative analyses. *Methods Ecol. Evol.* 3, 940-947.
2322 <http://dx.doi.org/10.1111/j.2041-210X.2012.00220.x>.

2323 Friedman, M., Keck, B. P., Dornburg, A., Eytan R. I., Martin, C. H., Hulsey, C. D., Wainwright, P. C., Near, T. J.,
2324 2013. Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting. *Proc. R.*
2325 *Soc. Lond. B* 280, 20131733. <http://dx.doi.org/10.1098/rspb.2013.1733>.

2326 Fritz, S. C., Cumming, B. F., Gasse, F., Laird, K. R., 1999. Numerical methods for the analysis of diatom
2327 assemblage data. In: Stoermer, E. F., Smol, J. P. (Eds.), *The Diatoms: Applications for the Environmental and*
2328 *Earth Sciences*. Cambridge University Press, Cambridge, pp. 41-72.

2329 García, N., López-Elías, J. A., Miranda, A., Martínez-Porchas, M., Huerta, N., García, A., 2012. Effect of salinity
2330 on growth and chemical composition of the diatom *Thalassiosira weissflogii* at three culture phases. *Lat. Am. J.*
2331 *Aquat. Res.* 40, 435-440.

2332 Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D., Drummond, A. J., 2015. Bayesian total
2333 evidence dating reveals the recent crown radiation of penguins. *ArXiv150604797 Q-Bio*.

2334 Genner, M. J., Knight, M. E., Haesler, M. P., Turner, G. F., 2010. Establishment and expansion of Lake Malawi
2335 rock fish populations after a dramatic Late Pleistocene lake level rise. *Mol. Ecol.* 19, 170-182.

2336 Genner, M. J., Nichols, P., Carvalho, G. R., Robinson, R. L., Shaw, P. W., Smith, A., Turner, G. F., 2007.
2337 Evolution of a cichlid fish in a Lake Malawi satellite lake. *Proc. Biol. Sci.* 274, 2249-2257.

2338 Ghinassi, M., D'oriano, F., Benvenuti, M., Fedi, M., Awramik, S., 2015. Lacustrine facies in response to
2339 millennial-century-scale climate changes (Lake Hayk, northern Ethiopia). *J. Sediment. Res.* 85, 381-398.

2340 Glombitza, C., Stockhecke, M., Schubert, C. J., Vetter, A., Kallmeyer, J., 2013. Sulfate reduction controlled by
2341 organic matter availability in deep sediment cores from the saline, alkaline Lake Van (Eastern Anatolia,
2342 Turkey). *Front. Microbiol.* 4, 209.

2343 Gonzalez-Voyer, A., Kolm, N., 2011. Rates of phenotypic evolution of ecological characters and sexual traits
2344 during the Tanganyikan cichlid adaptive radiation. *J. Evol. Biol.* 24, 2378-2388.
2345 <http://dx.doi.org/10.1111/j.1420-9101.2011.02365.x>.

- 2346 Goodwin, D. H., Schöne, B. R., Dettman, D. L., 2003. Resolution and fidelity of oxygen isotopes as
2347 paleotemperature proxies in bivalve mollusk shells: models and observations. *Palaios* 18, 110-125.
- 2348 Gorthner, A., 1994. What is an ancient lake? In: Martens, K., Goddeeris, B., Coulter, G. (Eds.), *Speciation in*
2349 *Ancient Lakes*. Arch. Hydrobiol. 44, 97-100.
- 2350 Gross, M., 2012. The mysteries of the diatoms. *Curr. Biol.* 22, R581-R585.
2351 <http://dx.doi.org/10.1016/j.cub.2012.07.041>.
- 2352 Guru, D. S., Siddesha, S., Manjunath, S. (Eds.), 2013. *Texture in Classification of Pollen Grain Images*. Multimedia
2353 *Processing, Communication and Computing Applications, Lecture Notes in Electrical Engineering, Vol. 213*.
2354 Springer India, pp. 77-89.
- 2355 Hagelberg, E., Hofreiter, M., Keyser, C., 2015. Ancient DNA: the first three decades. *Philos. Trans. R. Soc. Lond.*
2356 *B* 370, 20130371.
- 2357 Hammarlund, D., Buchardt, B., 1996. Composite stable isotope records from a Late Weichselian lacustrine
2358 sequence at Graeenge, Lolland, Denmark: Evidence of Allerod and Younger Dryas environments. *Boreas* 25, 8-
2359 22.
- 2360 Hanisch, S., Ariztegui, D., Püttmann, W., 2003. The biomarker record of Lake Albano, central Italy - implications
2361 for Holocene aquatic system response to environmental change. *Org. Geochem.* 34, 1223-1235.
- 2362 Hansen, T. F., 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51, 1341-1351.
- 2363 Hansen, T. F., 2014. Use and misuse of comparative methods in the study of adaptation. In: Garamszegi, L. Z.
2364 (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology*. Springer
2365 Heidelberg, pp. 351-379.
- 2366 Hare, P. E., Abelson, P. H., 1968. Racemization of amino acids in fossil shells. *Yearbook of the Carnegie*
2367 *Institution of Washington* 66, 526-528.
- 2368 Harmon, L. J., Harrison, S., 2015. Species diversity is dynamic and unbounded at local and continental scales. *Am.*
2369 *Nat.* 185, 584-593. <http://dx.doi.org/10.1086/680859>.
- 2370 Harmon, L. J., Losos, J. B., Jonathan Davies, T., Gillespie, R. G., Gittleman, J. L., Bryan Jennings, W., Kozak, K.
2371 H., McPeck, M. A., Moreno-Roark, F., Near, T. J., Purvis, A., Ricklefs, R. E., Schluter, D., Schulte II, J. A.,
2372 Seehausen, O., Sidlauskas, B. L., Torres-Carvajal, O., Weir, J. T., Mooers, A. Ø., 2010. Early bursts of body
2373 size and shape evolution are rare in comparative data. *Evolution* 64, 2385-2396.
2374 <http://dx.doi.org/10.1111/j.1558-5646.2010.01025.x>
- 2375 Harmon, L. J., Schulte, J. A., Larson, A., Losos, J. B., 2003. Tempo and Mode of Evolutionary Radiation in

- 2376 Iguanian Lizards. *Science* 301, 961-964. <http://dx.doi.org/10.1126/science.1084786>.
- 2377 Harvey, P. H., May, R. M., Nee, S., 1994. Phylogenies without fossils. *Evolution* 48, 523.
- 2378 Harzhauser, M., Mandic, O., Kern, A. K., Piller, W. E., Neubauer, T. A., Albrecht, C., Wilke, T., 2013. Explosive
2379 demographic expansion by dreissenid bivalves as a possible result of astronomical forcing. *Biogeosciences* 10,
2380 8423-8431. <http://dx.doi.org/10.5194/bg-10-8423-2013>.
- 2381 Hauffe, T., Albrecht, C., Wilke, T., 2015. Gastropod diversification and community structuring processes in ancient
2382 Lake Ohrid: a metacommunity speciation perspective. *Biogeosci. Discuss.* 12, 16081-16103.
2383 <http://dx.doi.org/10.5194/bgd-12-16081-2015>.
- 2384 Heath, T. A., Huelsenbeck, J. P., Stadler, T., 2014. The fossilized birth–death process for coherent calibration of
2385 divergence-time estimates. *Proc. Natl. Acad. Sci. USA* 111, E2957-E2966.
2386 <http://dx.doi.org/10.1073/pnas.1319091111>.
- 2387 Hendy, E. J., Tomiak, P. J., Collins, M. J., Hellstrom, J., Tudhope, A. W., Lough, J. M., Penkman, K. E. H., 2012.
2388 Assessing amino acid racemization variability in coral intra-crystalline protein for geochronological
2389 applications. *Geochim. Cosmochim. Acta* 86, 338-353.
- 2390 Herder, F., Pfaender, J., Schliewen, U. K., 2008. Adaptive sympatric speciation of polychromatic “roundfin” sailfin
2391 silverside fish in Lake Matano (Sulawesi). *Evolution* 62, 2178-2195.
- 2392 Hipsley, C. A., Müller, J., 2014. Beyond fossil calibrations: realities of molecular clock practices in evolutionary
2393 biology. *Front. Genet.* 5, 138. <http://dx.doi.org/10.3389/fgene.2014.00138>.
- 2394 Ho, S. Y., 2014. The changing face of the molecular evolutionary clock. *Trends Ecol. Evol.* 29, 496-503.
2395 <http://dx.doi.org/10.1016/j.tree.2014.07.004>.
- 2396 Ho, S. Y., Duchêne, S., 2014. Molecular-clock methods for estimating evolutionary rates and timescales. *Mol. Ecol.*
2397 23, 5947-5965. <http://dx.doi.org/10.1111/mec.12953>.
- 2398 Ho, S. Y., Shapiro, B., 2011. Skyline-plot methods for estimating demographic history from nucleotide sequences.
2399 *Mol. Ecol. Resour.* 11, 423-434.
- 2400 Ho, S. Y., Tong, K. J., Foster, C. S., Ritchie, A. M., Lo, N., Crisp, M. D., 2015. Biogeographic calibrations for the
2401 molecular clock. *Biol. Lett.* 11, 20150194.
- 2402 Hofreiter, M., Serre, D., Poinar, H. N., Kuch, M., Pääbo, S., 2001. Ancient DNA. *Nat. Rev. Genet.* 2, 353-359.
- 2403 Holt, K. A., Bennett, K. D., 2014. Principles and methods for automated palynology. *New Phytol.* 203, 735-742.
- 2404 Holtvoeth, J., Rushworth, D., Imeri, A., Cara, M., Vogel, H., Wagner, T., Wolff, G. A., 2016. Improved end-
2405 member characterization of modern organic matter pools in the Ohrid Basin (Albania, Macedonia) and

2406 evaluation of new palaeoenvironmental proxies. *Biogeosci.*, 13, 795-816. <http://doi.org/10.5194/bg-13-795->
2407 2016

2408 Hooghiemstra, H., 1989. Quaternary and upper-Pliocene glaciations and forest development in the tropical Andes:
2409 evidence from a long high-resolution pollen record from the sedimentary basin of Bogota, Columbia.
2410 *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 72, 11-26.

2411 Houle, D., Mezey, J., Galpern, P., Carter, A., 2003. Automated measurement of *Drosophila* wings. *BMC Evol.*
2412 *Biol.* 3:25. <http://doi.org/10.1186/1471-2148-3-25>

2413 Huang, S., Roy, K., Jablonski, D., 2015. Origins, bottlenecks, and present-day diversity: Patterns of morphospace
2414 occupation in marine bivalves. *Evolution* 69, 735-746. <http://doi.org/10.1111/evo.12608>

2415 Huang, Y. S., Shuman, B., Wang, Y., Webb, T., 2002. Hydrogen isotope ratios of palmitic acid in lacustrine
2416 sediments record late quaternary climate variations. *Geology* 30, 1103-1106.

2417 Huang, Y. S., Shuman, B., Wang, Y., Webb, T., 2004. Hydrogen isotope ratios of individual lipids in lake
2418 sediments as novel tracers of climatic and environmental change: a surface sediment test. *J. Paleolimn.* 31, 363-
2419 375.

2420 Hubert, N., Calcagno, V., Etienne, R. S., Mouquet, N., 2015. Metacommunity speciation models and their
2421 implications for diversification theory. *Ecol. Lett.* 18, 864-881. <http://doi.org/10.1111/ele.12458>.

2422 Hunt, G., 2006. Fitting and comparing models of phyletic evolution: Random walks and beyond. *Paleobiology* 32,
2423 578-601.

2424 Hunt, G., 2012. paleoTS: Modeling evolution in paleontological time-series, version 0.4-4. Available at
2425 <http://cran.r-project.org/web/packages/paleoTS/index.html>. Accessed August, 2015.

2426 Hunt, G., Bell, M. A., Travis, M. P., 2008. Evolution toward a new adaptive optimum: phenotypic evolution in a
2427 fossil stickleback lineage. *Evolution* 62, 700-710.

2428 Inagaki, F., 2010. Deep seafloor microbial communities. In: *Encyclopedia of Life Sciences*. John Wiley & Sons,
2429 Ltd., Chichester, UK.

2430 Insinga, D. D., Tamburrino, S., Lirer, F., Vezzoli, L., Barra, M., De Lange, G. J., Tiepolo, M., Vallefucio, M.,
2431 Mazzola, S., Sprovieri, M., 2014. Tephrochronology of the astronomically-tuned KC01B deep-sea core, Ionian
2432 Sea: insights into the explosive activity of the Central Mediterranean area during the last 200 ka. *Quat. Sci. Rev.*
2433 85, 63-84. <http://dx.doi.org/10.1016/j.quascirev.2013.11.019>.

2434 Ishiwatari, R., Yamamoto, S., Shinoyama, S., 2006. Lignin and fatty acid records in Lake Baikal sediments over the
2435 last 130 kyr: A comparison with pollen records. *Org. Geochem.* 37, 1787-1802.

2436 Izart, A., Palhol, F., Gleixner, G., Elie, M., Blaise, T., Suarez-Ruiz, I., Sachsenhofer, R. F., Privalov, V. A., Panova,
2437 E. A., 2012. Palaeoclimate reconstruction from biomarker geochemistry and stable isotopes of *n*-alkanes from
2438 Carboniferous and Early Permian humic coals and limnic sediments in western and eastern Europe. *Org.*
2439 *Geochem.* 43, 125-149.

2440 Jalba, A. C., Wilkinson, M. H. F., Roerdink, J. B. T. M., Bayer, M. M., Juggins, S., 2005. Automatic diatom
2441 identification using contour analysis by morphological curvature scale spaces. *Mach. Vision Appl.* 16, 217-228.

2442 Jervey, M. T., 1988, Quantitative geological modeling of siliciclastic rock sequences and their seismic expression,
2443 In: Wilgus, C. K., Hastings, B. S., Kendall, C. G., Posamentier, H.W., Ross, C. A., Van Wagoner, J. C. (Eds.),
2444 *Sea Level Changes – an Integrated Approach*, vol 42. SEPM Spec. Publ., pp. 47-69.

2445 Jørgensen, T., Haile, J., Möller, P., Andreev, A., Boessenkool, S., Rasmussen, M., Kienast, F., Coissac, E.,
2446 Taberlet, P., Brochmann, C., Bigelow, N. H., Andersen, K., Orlando, L., Gilbert, M. T., Willerslev, E., 2012. A
2447 comparative study of ancient sedimentary DNA, pollen and microfossils from permafrost sediments of northern
2448 Siberia reveals long-term vegetational stability. *Mol. Ecol.* 21, 1989-2003.

2449 Jovanovska, E., Cvetkoska, A., Hauffe, T., Levkov, Z., Wagner, B., Sulpizio, R., Francke, A., Albrecht, C., Wilke,
2450 T., 2016. Differential resilience of ancient sister lakes Ohrid and Prespa to environmental disturbances during
2451 the Late Pleistocene. *Biogeosci.* 13, 1149-1161. <http://dx.doi.org/10.5194/bg-13-1149-2016>.

2452 Juggins, S., 2013. Quantitative reconstructions in palaeolimnology: new paradigm or sick science? *Quat. Sci. Rev.*
2453 64, 20-32. <http://doi.org/10.1016/j.quascirev.2012.12.014>

2454 Juggins, S., Birks, H. J. B., 2012. Quantitative environmental reconstructions from biological data. In: Birks, H. J.
2455 B., Lotter, A. F., Juggins, S., Smol, J. P. (Eds.), *Tracking environmental change using lake sediments: data*
2456 *handling and numerical techniques*, Vol. 5. Springer Netherlands, pp. 431-494.

2457 Kaandorp, R. J. G., Vonhof, H. B., Wesselingh, F. P., Pittman, L. R., Kroon, D., van Hinte, J. E., 2005. Seasonal
2458 Amazonian rainfall variation in the Miocene Climate Optimum. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 221,
2459 1-6.

2460 Kallmeyer, J., Grewe, S., Glombitza, C., Axel Kitte, J., 2015. Microbial abundance in lacustrine sediments: a case
2461 study from Lake Van, Turkey. *Int. J. Earth Sci. (Geol. Rundsch.)* 104, 1667-1677.

2462 Kallmeyer, J., Pockalny, R., Adhikari, R. R., Smith, D. C., D'Hondt, S., 2012. Global distribution of microbial
2463 abundance and biomass in subseafloor sediment. *Proc. Natl. Acad. Sci. USA* 109, 16213-16216.

2464 Kashiwaya, K., Ochiai, S., Sakai, H., Kawai, T., 2001. Orbit-related long-term climate cycles revealed in a 12-Myr
2465 continental record from Lake Baikal. *Nature* 410, 71-74.

2466 Kaufman, D. S., 2003a. Dating deep-lake sediments by using amino acid racemization in fossil ostracodes. *Geology*
2467 31, 1049-1052.

2468 Kaufman, D. S., 2003b. Amino acid palaeothermometry of Quaternary ostracodes from the Bonneville Basin, Utah.
2469 *Quat. Sci. Rev.* 22, 899-914.

2470 Kelts, K., Talbot, M. R., 1990. Lacustrine carbonates as geochemical archives of environmental change and biotic /
2471 abiotic interactions. In: Tilzer, M. M., Serruya, C. (Eds.), *Ecological structure and function in large lakes*.
2472 Springer, Heidelberg, pp. 288-315.

2473 Kenyon, W. E., 1997. Petrophysical Principles of Applications of NMR Logging. *The Log Analyst* 38, 21-43.

2474 Kermarrec, L., Bouchez, A., Rimet, R., Humbert, J.-F., 2013. First evidence of the existence of semi-cryptic species
2475 and of a phylogeographic structure in the *Gomphonema parvulum* (Kützing) Kützing complex. *Protist* 164, 686-
2476 705.

2477 Khursevich, G., 2006. Evolution of the extinct genera belonged to the family Stephanodiscaceae (Bacillariophyta)
2478 during the last eight million years in Lake Baikal. In: Ognjanova-Rumenova, N., Manoylov K. (Eds.), *Advances*
2479 *in phycological studies*. Pensoft Publishers, Sofia, pp. 73-89.

2480 Khursevich, G., Prokopenko, A. A., 2009. Diatom record from Lake Hovsgol, Mongolia, during the last 1 Ma: The
2481 results from the HDP-04 drill core. *Quatern. Int.* 205, 84-97.

2482 Kim, S. T., O'Neil, J. R., 1997. Equilibrium and nonequilibrium oxygen isotope effects in synthetic carbonates.
2483 *Geochim. Cosmochim. Acta* 61, 3461-3475.

2484 King, B., Lee, M. S. Y., 2015. Ancestral State Reconstruction, Rate Heterogeneity, and the Evolution of Reptile
2485 Viviparity. *Syst. Biol.* 64, 532-544. <http://dx.doi.org/10.1093/sysbio/syv005>.

2486 Kloster, M., Kauer, G., Beszteri, B., 2014. SHERPA: an image segmentation and outline feature extraction tool for
2487 diatoms and other objects. *BMC Bioinformatics* 15, 218. <http://dx.doi.org/10.1186/1471-2105-15-218>.

2488 Koblmüller, S., Odhiambo, E. A., Sinyinza, D., Sturmbauer, C., Sefc, K. M., 2015. Big fish, little divergence:
2489 phylogeography of Lake Tanganyika's giant cichlid, *Boulengerochromis microlepis*. *Hydrobiologia* 748, 29-38.
2490 <http://doi.org/10.1007/s10750-014-1863-z>.

2491 Koblmüller, S., Salzburger, W., Sturmbauer, C., 2004. Evolutionary relationships in the sand-dwelling cichlid
2492 lineage of Lake Tanganyika suggest multiple colonization of rocky habitats and convergent origin of biparental
2493 mouthbrooding. *J. Mol. Evol.* 58, 79-96.

2494 Koblmüller, S., Salzburger, W., Obermüller, B., Eigner, E., Sturmbauer, C., Sefc, K. M., 2011. Separated by sand,
2495 fused by dropping water: habitat barriers and fluctuating water levels steer the evolution of rock-dwelling

- 2496 cichlid populations in Lake Tanganyika. *Mol. Ecol.* 20, 2272-2290.
- 2497 Kocher, T. D., 2004. Adaptive evolution and explosive speciation: the cichlid fish model. *Nat. Rev. Genet.* 5, 288-
2498 298. <http://dx.doi.org/10.1038/nrg1316>.
- 2499 Konhauser, K., 2007. *Introduction to Geomicrobiology*. Blackwell Science Ltd., Oxford.
- 2500 Kornilova, A., Rosell-Melé, O., 2003. Application of microwave-assisted extraction to the analysis of biomarker
2501 climate proxies in marine sediments. *Org. Geochem.* 34, 1517-1523.
- 2502 Koskinen, M. T., Knizhin, I., Primmer, C. R., Schlötterer, C., Weiss, S., 2002. Mitochondrial and nuclear DNA
2503 phylogeography of *Thymallus* spp. (grayling) provides evidence of ice-age mediated environmental
2504 perturbations in the world's oldest body of fresh water, Lake Baikal. *Mol. Ecol.* 11, 2599-2611.
- 2505 Krause, J., 2010. From Genes to Genomes: What is New in Ancient DNA? *Mitteilungen der Gesell. für*
2506 *Urgeschichte* 19, 11-33.
- 2507 Kroll, O., Hershler, R., Albrecht, C., Terrazas, E. M., Apaza, R., Fuentealba, C., Wolff, C., Wilke, T., 2012. The
2508 endemic gastropod fauna of Lake Titicaca: correlation between molecular evolution and hydrographic history.
2509 *Ecol. Evol.* 2, 1517-1530. <http://dx.doi.org/10.1002/ece3.280>.
- 2510 Kröpelin, S., Verschuren, D., Lézine, A.-M., Eggermont, H., Cocquyt, C., Francus, P., Cazet, J.-P., Fagot, M.,
2511 Rumes, B., Russell, J. M., Darius, F., Conley, D. J., Schuster, M., Von Suchodoletz, H., Engstrom, D. R., 2008.
2512 Climate-driven ecosystem succession in the Sahara: the past 6000 years. *Science* 320, 765-768.
- 2513 Kulikovskiy, M., Lange-Bertalot, H., Metzeltin, D., Witkowski, A., 2012. Lake Baikal: Hotspot of endemic
2514 diatoms, I. *Iconogr. Diatomol.* 23, 1-607.
- 2515 Kylander, M. E., Bindler, R., Martínez Cortizas, A., Gallagher, K., Mörth, C.-M., Rauch, S., 2013. A Novel
2516 Geochemical Approach to Paleorecords of Dust Deposition and Effective Humidity: 8500 Years of Peat
2517 Accumulation at Store Mosse ("Big Bog"), Sweden. *Quat. Sci. Rev.* 69, 69-82.
- 2518 Lacey, J. H., Francke, A., Leng, M. J., Vane, C. H., Wagner, B., 2015. A high-resolution Late Glacial to Holocene
2519 record of environmental change in the Mediterranean from Lake Ohrid (Macedonia/Albania). *Int. J. Earth Sci.*
2520 104, 1623-1638.
- 2521 Lacey, J. H., Leng, M. J., Francke, A., Sloane, H. J., Milodowski, A., Vogel, H., Baumgarten, H., Zanchetta, G.,
2522 Wagner, B., 2016. Northern Mediterranean climate since the Middle Pleistocene: a 637 ka stable isotope record
2523 from Lake Ohrid (Albania/Macedonia). *Biogeosciences* 13, 1801-1820. [http://dx.doi.org/10.5194/bg-13-1801-](http://dx.doi.org/10.5194/bg-13-1801-2016)
2524 2016.
- 2525 Laj, C., Channell, J. E. T., 2009. Geomagnetic excursions. In: Kono, M., Schubert, G. (Eds.), *Geomagnetism*.

- 2526 Elsevier, Amsterdam, pp. 373-416.
- 2527 Lamb, H., Kebede, S., Leng, M. J., Ricketts, D., Telford, R., Umer, M., 2002. Origin and stable isotope
2528 composition of aragonite laminae in an Ethiopian crater lake. In: Odada, E., Olago, D. (Eds.), *The East African
2529 Great Lakes Region: Limnology, Palaeoclimatology and Biodiversity*, Advances in Global Research Series,
2530 Kluwer publishers, Dordrecht.
- 2531 Langlet, D., Alleman, L. Y., Plisnier, P. D., Hughes, H., André, L., 2007. Manganese content records seasonal
2532 upwelling in Lake Tanganyika mussels. *Biogeosciences* 4, 195-203.
- 2533 Lawing, A. M., Matzke, N. J., 2014. Conservation paleobiology needs phylogenetic methods, *Ecography* 37, 1109-
2534 1122. <http://dx.doi.org/10.1111/ecog.00783>.
- 2535 Leicher, N., Zanchetta, G., Sulpizio, R., Giaccio, B., Wagner, B., Nomade, S., Francke, A., Del Carlo, P., 2015.
2536 First tephrostratigraphic results of the DEEP site record from Lake Ohrid, Macedonia, *Biogeosci. Discuss.* 12,
2537 15411-15460. <http://dx.doi.org/10.5194/bgd-12-15411-2015>.
- 2538 Leng, M. J., Barker, P. A., 2006. A review of the oxygen isotope composition of lacustrine diatom silica for
2539 palaeoclimate reconstruction. *Earth-Sci. Rev.* 75, 5-27.
- 2540 Leng, M. J., Henderson, A. C. G., 2013. Recent advances in isotopes as palaeolimnological proxies. *J. Paleolimnol.*
2541 49, 481-496.
- 2542 Leng, M. J., Lewis, J. J., 2014. Oxygen isotopes in Molluscan shell: Applications in environmental archaeology.
2543 *Environ. Archaeol.* <http://dx.doi.org/10.1179/1749631414Y.0000000048>.
- 2544 Leng, M. J., Marshall, J. D., 2004. Palaeoclimate interpretation of stable isotope data from lake sediment archives.
2545 *Quat. Sci. Rev.* 23, 811-831.
- 2546 Leng, M. J., Roberts, N., Reed, J. M., Sloane, H. J., 1999b. Late Quaternary climatic and limnological variations
2547 based on carbon and oxygen isotope data from authigenic and ostracod carbonate in the Konya Basin, Turkey. *J.
2548 Paleolimnol.* 22, 187-204.
- 2549 Levkov, Z., Krstic, S., Metzeltin, D., Nakov, T., 2007. Diatoms of Lakes Prespa and Ohrid (Macedonia). *Iconogr.
2550 Diatomol.* 16, 1-603.
- 2551 Levkov, Z., Williams, D. M., 2012. Checklist of diatoms (Bacillariophyta) from Lake Ohrid and Lake Prespa
2552 (Macedonia), and their watersheds. *Phytotaxa* 45, 1-76.
- 2553 Liang, L., Xu, B., Chen, Y., Liu, Y., Cao, W., Fang, L., Feng, L., Goodchild, M. F., Gong, P., 2010. Combining
2554 spatial-temporal and phylogenetic analysis approaches for improved understanding on global H5N1
2555 transmission. *PLoS ONE* 5, e13575. <http://dx.doi.org/10.1371/journal.pone.0013575>.

- 2556 Lindhorst, K., Krastel, S., Reicherter, K., Stipp, M., Wagner, B., Schwenk, T., 2015. Sedimentary and tectonic
2557 evolution of Lake Ohrid (Macedonia/Albania). *Basin Res.* 27, 84-101. <http://dx.doi.org/10.1111/bre.12063>.
- 2558 Lisiecki, L. E., Raymo, M. E., 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records.
2559 *Paleoceanography* 20, PA1003. <http://dx.doi.org/10.1029/2004pa001071>.
- 2560 Litt, T., Pickarski, N., Heumann, G., Stockhecke, M., Tzedakis, P. C., 2014. A 600,000 year long continental pollen
2561 record from Lake Van, eastern Anatolia (Turkey). *Quat. Sci. Rev.* 104, 30-41.
- 2562 Lovell, M., Parkinson, N. (Eds.), 2002. *Geological Applications of Well Logs*. AAPG Methods in Exploration
2563 Series No. 13. Tulsa.
- 2564 Lowe, J., Walker, M., 2015. Amino acid geochronology (Chapter 5, Section 5.6.1). In: *Reconstructing Quaternary*
2565 *Environments*, 3rd ed. Routledge, pp. 332-339.
- 2566 Luthi, S. M., 2001. *Geological well logs, their use in reservoir modelling*. Springer, Berlin.
- 2567 Lyons, R. P., Scholz, C. A., Cohen, A. S., King, J. W., Brown, E. T., Ivory, S. J., Johnson, T. C., Deino, A. L.,
2568 Reinthal, P. N., McGlue, M. M., Blome, M. W., 2015. Continuous 1.3-million-year record of East African
2569 hydroclimate, and implications for patterns of evolution and biodiversity. *Proc. Natl. Acad. Sci.* 112, 15568–
2570 15573. <http://dx.doi.org/10.1073/pnas.1512864112>
- 2571 Mackay, A. W., 2007. The paleoclimatology of Lake Baikal: a diatom synthesis and prospectus. *Earth-Sci. Rev.*
2572 82,181-215.
- 2573 Mackay, A. W., Battarbee, R. W., Flower, R. J., Granin, N. G., Jewson, D. H., Ryves, D. B., Sturm, M., 2003.
2574 Assessing the potential for developing internal diatom-based inference models in Lake Baikal. *Limnol.*
2575 *Oceanogr.* 48, 1183-1192. <http://dx.doi.org/10.4319/lo.2003.48.3.1183>.
- 2576 Mackay, A. W., Edlund, M. B., Khursevich, G., 2010. Diatoms in ancient lakes. In: Smol, J. P., Stoermer, E. F.
2577 (Eds.), *The Diatoms: Applications for the Environmental and Earth Sciences*, 2nd ed. Cambridge University
2578 Press, Cambridge, pp. 209-228. <http://dx.doi.org/10.1017/CBO9780511763175.012>.
- 2579 Magee, J. W., Bowler, J. M., Miller, G. H., Williams, D., 1995. Stratigraphy, sedimentology, chronology and
2580 palaeohydrology of Quaternary lacustrine deposits at Madigan Gulf, Lake Eyre, South Australia. *Palaeogeogr.*
2581 *Palaeoclimatol. Palaeoecol.* 113, 3-42.
- 2582 Magyar, I., Müller, P. M., Sztanó, O., Babinszki, E., Lantos, M., 2006. Oxygen-related facies in Lake Pannon
2583 deposits (Upper Miocene) at Budapest-Köbánya. *Facies*, 52, 209-220.
- 2584 Mahler, D. L., Ingram, T., Revell, L. J., Losos, J. B., 2013. Exceptional Convergence on the Macroevolutionary
2585 Landscape in Island Lizard Radiations. *Science* 341, 292-295. <http://dx.doi.org/10.1126/science.1232392>.

- 2586 Mansilla, C., Novais, M. H., Faber, E., Martínez-Martínez, D., De Hosson, J. T., 2015. On the 3D reconstruction of
2587 diatom frustules: a novel method, applications, and limitations. *J. Appl. Phycol.* 27, 1-14.
- 2588 Marcos, J. V., Nava, R., Cristóbal, G., Redondo, R., Escalante-Ramírez, B., Bueno, G., Déniz, Ó., González-Porto
2589 A., Pardo, C., Chung, F., Rodríguez, T., 2015. Automated pollen identification using microscopic imaging and
2590 texture analysis. *Micron* 68, 36-46. <http://dx.doi.org/10.1016/j.micron.2014.09.002>.
- 2591 Mardis, E. R., 2011. A decade's perspective on DNA sequencing technology. *Nature* 470, 198-203.
2592 <http://dx.doi.org/10.1038/nature09796>.
- 2593 Martens, K., 1997. Speciation in ancient lakes. *Trends Ecol. Evol.* 12, 177-182.
- 2594 Marynowski, L., Rakociński, M., Borcuch, E., Kremer, B., Schubert, B. A., Jahren, A. H., 2011. Molecular and
2595 petrographic indicators of redox conditions and bacterial communities after the F/F mass extinction (Kowala,
2596 Holy Cross Mountains, Poland), *Palaeogeogr. Palaeoclim. Palaeoecol.* 306, 1-14.
- 2597 Mazzini, I., Gliozzi, E., Koci, R., Soulie-Märsche, I., Zanchetta, G., Baneschi, I., Sadori, L., Giardini, M., Van
2598 Welden, A., Bushati, S., 2015. Historical evolution and Middle to Late Holocene environmental changes in
2599 Lake Shkodra (Albania): New evidences from micropaleontological analysis. *Palaeogeogr. Palaeoclim.*
2600 *Palaeoecol.* 419, 47-59.
- 2601 McCoy, W. D., 1987. Quaternary aminostratigraphy of the Bonneville Basin, western United States. *Bull. Geol.*
2602 *Soc. Am.* 98, 99-112.
- 2603 McCrea, J. M., 1950. On the isotopic chemistry of carbonates and palaeo-temperature scale. *J. Chem. Phys.* 18,
2604 849-857.
- 2605 Meixner, M. J., Lüter, C., Eckert, C., Itskovich, V., Janussen, D., von Rintelen, T., Hess, W. R., 2007. Phylogenetic
2606 analysis of freshwater sponges provide evidence for endemism and radiation in ancient lakes. *Mol. Phylogenet.*
2607 *Evol.* 45, 875-886.
- 2608 Melles, M., Brigham-Grette, J., Glushkova, O. Y., Minyuk, P. S., Nowaczyk, N. R., Hubberten, H. W., 2007.
2609 Sedimentary geochemistry of core PG1351 from Lake El'gygytyn - a sensitive record of climate variability in
2610 the East Siberian Arctic during the past three glacialinterglacial cycles. *J. Paleolimnol.* 37, 89-104.
- 2611 Melles, M., Brigham-Grette, J., Minyuk, P. S., Nowaczyk, N. R., Wennrich, V., DeConto, R. M., Anderson, P. M.,
2612 Andreev, A. A., Coletti, A., Cook, T. L., Haltia-Hovi, E., Kukkonen, M., Lozhkin, A. V., Rosén, P., Tarasov, P.,
2613 Vogel, H., Wagner, B., 2012. 2.8 Million years of Arctic Climate Change from Lake El'gygytyn, NE Russia.
2614 *Science* 337, 315-320.
- 2615 Merrill, R. T., McElhinny, M. W., McFadden, P. L., 2006. The magnetic field of the Earth paleomagnetism, the

2616 core, and the deep mantle (2nd ed). Elsevier, Burlington.

2617 Metzker, M. L., 2010. Sequencing technologies - the next generation. *Nat. Rev. Genet.* 11, 31-46.

2618 <http://dx.doi.org/10.1038/nrg2626>.

2619 Meyer-Jacob, C., Vogel, H., Boxberg, F., Rosén, P., Weber, M. E., Bindler, R., 2014a. Independent measurement
2620 of biogenic silica in sediments by FTIR spectroscopy and PLS regression. *J. Paleolimnol.* 52, 245-255.

2621 Meyer-Jacob, C., Vogel, H., Gebhardt, C., Wennrich, V., Melles, M., Rosén, P., 2014b. Biogeochemical variability
2622 during the past 3.6 million years recorded by FTIR spectroscopy in the sediment record of Lake El'gygytgyn,
2623 Far East Russian Arctic. *Clim. Past* 10, 209-220.

2624 Meyers, P. A., 2003. Applications of organic geochemistry to paleolimnological reconstructions: a summary of
2625 examples from the Laurentian Great Lakes. *Org. Geochem.* 34, 261-289.

2626 Miller, G. H., Hare, P. E., 1980. Amino acid geochronology: integrity of the carbonate matrix and potential of
2627 molluscan fossils. In: Hare, P. E., Hoering, T. C., King Jr., K. (Eds.), *Biogeochemistry of Amino Acids*. Wiley,
2628 New York, pp. 415-443.

2629 Mischke, S., Rajabov, I., Mustaeva, N., Zhang, C., Herzsuh, U., Boomer, I., Brown, E. T., Andersen, N., Myrbo,
2630 A., Ito, E., Schudack, M. E., 2010. Modern hydrology and late Holocene history of Lake Karakul, eastern
2631 Pamirs (Tajikistan): a reconnaissance study. *Palaeogeogr. Palaeoclim. Palaeoecol.* 289, 10-24.

2632 Morlon, H., 2014. Phylogenetic approaches for studying diversification. *Ecol. Lett.* 17, 508-525.

2633 <http://dx.doi.org/10.1111/ele.12251>.

2634 Morlon, H., Parsons, T. L., Plotkin, J. B., 2011. Reconciling molecular phylogenies with the fossil record. *Proc.*
2635 *Natl. Acad. Sci. USA* 108, 16327-16332. <http://dx.doi.org/10.1073/pnas.1102543108>.

2636 Mosleh, M. A., Manssor, H., Malek, S., Milow, P., Salleh, A., 2012. A preliminary study on automated freshwater
2637 algae recognition and classification system. *BMC Bioinformatics* 13, 1-13.

2638 Mourguiart, P., Corrège, T., Wirrmann, D., Argollo, J., Montenegro, M.E., Pourchet, M., Carbonel, P., 1998.
2639 Holocene palaeohydrology of Lake Titicaca estimated from an ostracod-based transfer function. *Palaeogeogr.*
2640 *Palaeoclimatol. Palaeoecol.* 143, 51-72. [http://dx.doi.org/10.1016/S0031-0182\(98\)00068-6](http://dx.doi.org/10.1016/S0031-0182(98)00068-6).

2641 Müller, J., Oberhänsli, H., Melles, M., Schwab, M., Rachold, V., Hubberten, H.-W., 2001. Late Pliocene
2642 sedimentation in Lake Baikal: implications for climatic and tectonic change in SE Siberia. *Palaeogeogr.*
2643 *Palaeoclimatol. Palaeoecol.* 174, 305-326. [http://dx.doi.org/10.1016/S0031-0182\(01\)00320-0](http://dx.doi.org/10.1016/S0031-0182(01)00320-0).

2644 Munro, M. A. R., Kresler, A. M., Battarbee, R. W., Juggins, S., Stevenson, A. C., Anderson, D.S., Anderson, N. J.,
2645 Berge, F., Birks, H. J. B., Davis, R. B., Fritz, S. C., Haworth, E. Y., Jones, V. J., Kingston, J. C., Rengberg, I.,

2646 1990. Diatom quality control and data handling. *Philos. Trans. R. Soc. Lond. B* 327, 257-261.

2647 Muschick, M., Indermaur, A., Salzburger, W., 2012. Convergent evolution within an adaptive radiation of cichlid
2648 fishes. *Curr. Biol.* 22, 2362-2368. <http://dx.doi.org/10.1016/j.cub.2012.10.048>.

2649 Muschick, M., Nosil, P., Roesti, M., Dittmann, M. T., Harmon, L., Salzburger, W., 2014. Testing the stages model
2650 in the adaptive radiation of cichlid fishes in East African Lake Tanganyika. *Proc. R. Soc. Lond. B* 281,
2651 20140605. <http://dx.doi.org/10.1098/rspb.2014.0605>.

2652 Naeher, S., Gilli, A., North, R., Hamann, Y., Schubert, C., 2013. Tracing bottom water oxygenation with
2653 sedimentary Mn/Fe ratios in Lake Zurich, Switzerland. *Chem. Geol.* 352 125-133.

2654 Neubauer, T. A., Harzhauser, M., Georgopoulou, E., Kroh, A., Mandic, O., 2015. Tectonics, climate, and the rise
2655 and demise of continental aquatic species richness hotspots. *Proc. Natl. Acad. Sci. USA* 112, 11478-11483.
2656 <http://dx.doi.org/10.1073/pnas.1503992112>.

2657 Nevado, B., Mautner, S., Sturmbauer, C., Verheyen, E., 2013. Water-level fluctuations and metapopulation
2658 dynamics as drivers of genetic diversity in populations of three Tanganyikan cichlid fish species. *Mol. Ecol.* 22,
2659 3933-3948.

2660 Ng, J., Smith, S. D., 2014. How traits shape trees: new approaches for detecting character state-dependent lineage
2661 diversification. *J. Evol. Biol.* 27, 2035-2045. <http://dx.doi.org/10.1111/jeb.12460>.

2662 Niemi, T. M., Ben-Avraham, Z., Gat, J. R., 1997. The Dead Sea, the lake and its Setting, *Monographs on Geology
2663 and Geophysics* No. 36. Oxford University Press, New York.

2664 Nowaczyk, N. R., Haltia-Hovi, E. M., Ulbricht, D., Wennrich, R., Sauerbrey, M. A., Rosén, P., Vogel, H., Francke,
2665 A., Meyer-Jacob, C., Andreev, A. A., Lozhkin, A., 2013. Chronology of Lake El'gygytyn sediments – a
2666 combined magnetostratigraphic, palaeoclimatic and orbital tuning study based on multi-parameter analyses.
2667 *Clim. Past* 9, 2413-2432. <http://dx.doi.org/10.5194/cp-9-2413-2013>.

2668 O'Meara, B. C., Ané, C., Sanderson, M. J., Wainwright, P. C., 2006. Testing for different rates of continuous trait
2669 evolution using likelihood. *Evolution* 60, 922-933. <http://dx.doi.org/10.1111/j.0014-3820.2006.tb01171.x>

2670 Okuda, N., Watanabe, K., Fukumori, K., Nakano, S. I., Nakazawa, T., 2014. Origin and diversification of
2671 freshwater fishes in Lake Biwa. In: Okuda, N., Watanabe, K., Fukumori, K., Nakano, S. I., Nakazawa, T. (Eds.),
2672 *Biodiversity in aquatic systems and environments*. Springer Japan, pp. 1-19.

2673 Ortiz, J. E., Torres, T., Delgado, A., Julià, R., Llamas, F. J., Soler, V., Delgado, J., 2004. Numerical dating
2674 algorithms of amino acid racemization ratios analyzed in continental ostracodes of the Iberian Peninsula
2675 (Spain). Application to Guadix-Baza Basin (southern Spain). *Quat. Sci. Rev.* 23, 717-730.

2676 O'Sullivan, P., 2004. Palaeolimnology. In: O'Sullivan, P. E., Reynolds, C. S. (Eds.), The Lakes Handbook. Volume
2677 1: Limnology and Limnetic Ecology. Blackwell Publishing, Oxford, pp. 609-666.

2678 Ostrom, P. H., Schall, M., Gandhi, H., Shen, T.-L., Hauschka, P. V., Strahler, J. R., Gage, D. A., 2000. New
2679 strategies for characterizing ancient proteins using matrix-assisted laser desorption ionization mass
2680 spectrometry. *Geochim. Cosmochim. Acta* 64, 1043-1050.

2681 Oviatt, C. G., Thompson, R. S., Kaufman, D. S., Bright, J., Forester, R. M., 1999. Reinterpretation of the Burmester
2682 Core, Bonneville Basin, Utah. *Quat. Res.* 52, 180-184.

2683 Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L.,
2684 Hofreiter, M., 2004. Genetic Analyses from Ancient DNA. *Annu. Rev. Genet.* 38, 645-679.

2685 Pansu, J., Giguet-Covex, C., Ficetola, G. F., Gielly, L., Boyer, F., Zinger, L., Arnaud, F., Poulénard, J., Taberlet, P.,
2686 Choler, P., 2015. Reconstructing long-term human impacts on plant communities: an ecological approach based
2687 on lake sediment DNA. *Mol. Ecol.* 24, 1485-1498. <http://dx.doi.org/10.1111/mec.13136>.

2688 Parducci, L., Matetovici, I., Fontana, S. L., Bennett, K. D., Suyama, Y., Haile, J., Kjær, K. H., Larsen, N. K.,
2689 Drouzas, A. D., Willerslev, E., 2013. Molecular- and pollen-based vegetation analysis in lake sediments from
2690 central Scandinavia. *Mol. Ecol.* 22, 3511-3524. <http://dx.doi.org/10.1111/mec.12298>.

2691 Parducci, L., Väliiranta, M., Sakari Salonen, J., Ronkainen, T., Matetovici, I., Fontana, S. L., Eskola, T., Sarala, P.,
2692 Suyama, Y., 2015. Proxy comparison in ancient peat sediments: pollen, microfossil and plant DNA. *Philos.*
2693 *Trans. R. Soc. Lond. B* 370. online first. <http://dx.doi.org/10.1098/rstb.2013.0382>.

2694 Paulissen, E., Luthi, S. M., Grunert, P., Ćorić, S., Harzhauser, M., 2011. Integrated high-resolution stratigraphy of a
2695 Middle to Late Miocene sedimentary sequence in the central part of the Vienna Basin. *Geol. Carp.* 62, 155-169.

2696 Pawlowska, M. M., Butterfield, N. J., Brocks, J. J., 2013. Lipid taphonomy in the Proterozoic and the effect of
2697 microbial mats on biomarker preservation. *Geology* 41, 103-106.

2698 Pearson, P. N., Ezard, T. H. G., 2014. Evolution and speciation in the Eocene planktonic foraminifer *Turborotalia*.
2699 *Paleobiology* 40, 130-143.

2700 Peck, J. A., King, J. W., Colman, S. M., Kravchinsky, V. A., 1996. An 84-kyr paleomagnetic record from the
2701 sediments of Lake Baikal, Siberia. *J. Geophys. Res.* 101, 11365-11385.

2702 Pedersen, M. W., Ginolhac, A., Orlando, L., Olsen, J., Andersen, K., Holm, J., Funder, S., Willerslev, E., Kjær, K.
2703 H., 2013. A comparative study of ancient environmental DNA to pollen and microfossils from lake sediments
2704 reveals taxonomic overlap and additional plant taxa. *Quat. Sci. Rev.* 75, 161-168.
2705 <http://dx.doi.org/10.1016/j.quascirev.2013.06.006>.

2706 Penkman, K. E. H., Kaufman, D. S., Maddy, D., Collins, M. J., 2008. Closed-system behaviour of the intra-
2707 crystalline fraction of amino acids in mollusc shells. *Quat. Geochronol.* 3, 2-25.

2708 Penkman, K. E. H., Preece, R. C., Bridgland, D. R., Keen, D. H., Meijer, T., Parfitt, S. A., White, T. S., Collins, M.
2709 J., 2011. A chronological framework for the British Quaternary based on *Bithynia* opercula. *Nature* 476, 446-
2710 449.

2711 Penkman, K. E. H., Preece, R. C., Bridgland, D. R., Keen, D. H., Meijer, T., Parfitt, S. A., White, T. S., Collins, M.
2712 J., 2013. An aminostratigraphy for the British Quaternary based on *Bithynia* opercula. *Quat. Sci. Rev.* 61, 111-
2713 134.

2714 Pepe, C., Giardini, M., Giraudi, C., Masi, A., Mazzini, I., Sadori, L., 2013. Plant landscape and environmental
2715 changes recorded in marginal marine environments: the ancient Roman harbour of Portus (Rome, Italy). *Quat.*
2716 *Int.* 303, 73-81.

2717 Prokopenko, A. A., Hinnov, L. A., Williams, D. F., Kuzmin, M. I., 2006. Orbital forcing of continental climate
2718 during the Pleistocene: a complete astronomically tuned climatic record from Lake Baikal, SE Siberia. *Quat.*
2719 *Sci. Rev.* 25, 3431-3457. <http://dx.doi.org/10.1016/j.quascirev.2006.10.002>.

2720 Quental, T. B., Marshall, C. R., 2010. Diversity dynamics: molecular phylogenies need the fossil record. *Trends*
2721 *Ecol. Evol.* 25, 434-441.

2722 R Development Core Team, 2015. R: A language and environment for statistical computing (R Foundation for
2723 Statistical Computing). Available at <http://www.R-project.org>, v3.3.1. Accessed June 20, 2015.

2724 Rabosky, D. L., 2010. Extinction rates should not be estimated from molecular phylogenies. *Evolution* 64, 1816-
2725 1824. <http://dx.doi.org/10.1111/j.1558-5646.2009.00926.x>

2726 Rabosky, D. L., 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on
2727 phylogenetic trees. *PLoS ONE* 9, e89543. <http://dx.doi.org/10.1371/journal.pone.0089543>.

2728 Rabosky, D. L., Goldberg, E. E., 2015. Model Inadequacy and Mistaken Inferences of Trait-Dependent Speciation.
2729 *Syst. Biol.* 64, 340-355. <http://dx.doi.org/10.1093/sysbio/syu131>.

2730 Rabosky, D. L., Hurlbert, A. H., 2015. Species richness at continental scales is dominated by ecological limits. *Am.*
2731 *Nat.* 185, 572-583. <http://dx.doi.org/10.1086/680850>.

2732 Rabosky, D. L., McCune, A. R., 2010. Reinventing species selection with molecular phylogenies. *Trends Ecol.*
2733 *Evol.* 25, 68-74. <http://dx.doi.org/10.1016/j.tree.2009.07.002>.

2734 Recasens, C., Ariztegui D., Maidana, N. I., Zolitschka, B., PASADO Scientific Team, 2015. Diatoms as indicators
2735 of hydrological and climatic changes in Laguna Potrok Aike (Patagonia) since the Late Pleistocene.

2736 Palaeogeogr. Palaeoclim. Palaeoecol. 417, 309–319. <http://dx.doi.org/10.1016/j.palaeo.2014.09.021>.

2737 Reed, J. M., Cvetkoska, A., Levkov, Z., Vogel, H., Wagner, B., 2010. The last glacial-interglacial cycle in Lake
2738 Ohrid (Macedonia/Albania): testing diatom response to climate. *Biogeosciences* 7, 3083-3094.

2739 Reinthal, P. N., Cohen, A. S., Dettman, D. L., 2011. Fish fossils as paleo-indicators of ichthyofauna composition
2740 and climatic change in Lake Malawi, Africa. *Palaeogeogr. Palaeoclim. Palaeoecol.* 303, 126-132.

2741 Rethemeyer, J., Schubotz, F., Talbot, H. M., Cooke, M. P., Hinrichs, K.-U., Mollenhauer, G., 2010. Distribution of
2742 polar membrane lipids in permafrost soils and sediments of a small high Arctic catchment. *Org. Geochem.* 41,
2743 1130-1145.

2744 Revell, L. J., 2012. *Phytools: an R package for phylogenetic comparative biology (and other things)*. *Methods Ecol.*
2745 *Evol.* 3, 217-223. <http://dx.doi.org/10.1111/j.2041-210X.2011.00169.x>.

2746 Richter, G., Wedmann, S., 2005. Ecology of the Eocene Lake Messel revealed by analysis of small fish coprolites
2747 and sediments from a drilling core. *Palaeogeogr. Palaeoclim. Palaeoecol.* 223, 147-161.

2748 Rider, M., Kennedy, M., 2011. *The Geological Interpretation of Well Logs*, 3rd ed. Rider-French Consulting
2749 Limited.

2750 Rioual, P., Mackay, A. W., 2005. A diatom record of centennial resolution for the Kazantsevo Interglacial stage in
2751 Lake Baikal (Siberia). *Glob. Planet. Chang.* 46, 199-219.

2752 Ripley, B. D., 1976. The second-order analysis of stationary point processes. *J. Appl. Probab.* 13, 255-266.
2753 <http://dx.doi.org/10.2307/3212829>.

2754 Roberts, A. P., Tauxe, L., Heslop, D., 2013. Magnetic paleointensity stratigraphy and high-resolution Quaternary
2755 geochronology: successes and future challenges. *Quat. Sci. Rev.* 61, 1-16.
2756 <http://dx.doi.org/10.1016/j.quascirev.2012.10.036>.

2757 Roberts, A. P., Winklhofer, M., 2004. Why are geomagnetic excursions not always recorded in sediments?
2758 Constraints from post-depositional remanent magnetization lock-in modelling. *Earth Planet. Sci. Lett.* 227, 345-
2759 359.

2760 Rogers, A. R., Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic
2761 differences. *Mol. Biol. Evol.* 9, 552-569.

2762 Rohrssen, M., Love, G. D., Fischer, W., Finnegan, S., Fike, D. A., 2013. Lipid biomarkers record fundamental
2763 changes in the microbial community structure of tropical seas during the Late Ordovician Hirnantian glaciation.
2764 *Geology* 41, 127-130.

2765 Rose, D. T., Cox, E. J., 2014. What constitutes *Gomphonema parvulum*? Long-term culture studies show that some

- 2766 varieties of *G. parvulum* belong with other *Gomphonema* species. *Plant Ecol. Evol.* 147, 366-373.
- 2767 Round, F. E., Crawford, R., Mann, D. G., 1990. *The Diatoms. Morphology and biology of the genera.* Cambridge
2768 University Press, Cambridge, 747 pp.
- 2769 Røy, H., Kallmeyer, J., Adhikari, R. R., Pockalny, R., Jørgensen, B. B., D'Hondt, S., 2012. Aerobic microbial
2770 respiration in 86-million-year-old deep-sea red clay. *Science* 336, 922-925.
- 2771 Russell, J. M., Vogel, H., Konecky, B. L., Bijaksana, S., Huang, Y., Melles, M., Wattrus, N., Costa, K., King, J. W.,
2772 2014. Glacial forcing of central Indonesian hydroclimate since 60,000 y B.P. *Proc. Natl. Acad. Sci. USA* 111,
2773 5100-5105.
- 2774 Russell, J., Bijaksana, S., 2012. The Towuti Drilling Project: paleoenvironments, biological evolution, and
2775 geomicrobiology of a tropical Pacific lake. *Sci. Drill.* 14, 68-71.
- 2776 Ryves, D. B., Battarbee, R. W., Juggins, S., Fritz, S. C., Anderson, N. J., 2006. Physical and chemical predictors of
2777 diatom dissolution in freshwater and saline lake sediments in North America and West Greenland. *Limnol.*
2778 *Oceanogr.* 51, 1355-1368.
- 2779 Sachse, D., Billault, I., Bowen, G. J., Chikaraishi, Y., Dawson, T. E., Feakins, S. J., Freeman, K. H., Magill, C. R.,
2780 McInerney, F. A., van der Meer, M. T. J., Polissar, P. J., Robins, R. J., Sachs, J. P., Schmidt, H.-L., Sessions, A.
2781 L., White, J. W. C., West, J. B., Kahmen, A., 2012. Molecular paleohydrology: interpreting the hydrogen-
2782 isotopic composition of lipid biomarkers from photosynthesizing organisms. *Annu. Rev. Earth Planet. Sci.* 40,
2783 221-249.
- 2784 Sachse, D., Radke, J., Gleixner, G., 2004. Hydrogen isotope ratios of recent lacustrine sedimentary *n*-alkanes record
2785 modern climate variability. *Geochim. Cosmochim. Acta* 68, 4877-4889.
- 2786 Sadori, L., Koutsodendris, A., Masi, A., Bertini, A., Combourieu-Nebout, N., Francke, A., Kouli, K., Joannin, S.,
2787 Mercuri, A. M., Panagiotopoulos, K., Peyron, O., Torri, P., Wagner, B., Zanchetta, G., Donders, T. H., 2016.
2788 Pollen-based paleoenvironmental and paleoclimatic change at Lake Ohrid (south-eastern Europe) during the
2789 past 500 ka. *Biogeosci.* 13, 1423-1437. <http://dx.doi.org/10.5194/bg-13-1423-2016>.
- 2790 Sadori, L., Mercuri, A. M., Mariotti Lippi, M., 2010. Reconstructing past cultural landscape and human impact
2791 using pollen and plant macroremains. *Plant Biology* 144, 940-951.
2792 <http://dx.doi.org/10.1080/11263504.2010.491982>.
- 2793 Salzburger, W., Van Bocxlaer, B., Cohen, A. S., 2014. Ecology and evolution of the African Great Lakes and their
2794 faunas. *Annu. Rev. Ecol. Evol. Syst.* 45, 519-545. <http://dx.doi.org/10.1146/annurev-ecolsys-120213-091804>.
- 2795 Saros, J. E., Stone, J. R., Pederson, G.T., Slemmons, K. E. H., Spanbauer, T., Schliep, A., Cahl, D., Williamson, C.

2796 E., Engstrom, D. R., 2012. Climate-induced changes in lake ecosystem structure inferred from coupled neo- and
2797 paleoecological approaches. *Ecology* 93, 2155-2164.

2798 Schluter, D., 2000. *The ecology of adaptive radiation*, Oxford Series in Ecology and Evolution. Oxford University
2799 Press, Oxford.

2800 Schmieder, R., Edwards, R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27,
2801 863-864.

2802 Scholz, C. A., Klitgord, K. D., Hutchinson, D. R., Ten Brink, U. S., Zonenshain, L. P., Golmshtok, A. Y., Moore,
2803 T. C., 1993. Results of 1992 seismic reflection experiment in Lake Baikal. *Eos, Trans. Amer. Geophys. Union*
2804 74, 465-470.

2805 Scholz, C. A., Cohen, A. S., Johnson, T. C., King, J., Talbot, M. R., Brown, E. T., 2011. Scientific drilling in the
2806 Great Rift Valley: the 2005 Lake Malawi Scientific Drilling Project – an overview of the past 145,000 years of
2807 climate variability in Southern Hemisphere East Africa. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 303, 3-19.
2808 <http://dx.doi.org/10.1016/j.palaeo.2010.10.030>.

2809 Schön, I., Martens, K., 2004. Adaptive, pre-adaptive and non-adaptive components of radiations in ancient lakes: a
2810 review. *Org. Divers. Evol.* 4, 137-156. <http://dx.doi.org/10.1016/j.ode.2004.03.001>.

2811 Schön, I., Martens, K., 2011. Molecular analyses of ostracod flocks from Lake Baikal and Lake Tanganyika.
2812 *Hydrobiologia* 682, 91-110. <http://dx.doi.org/10.1007/s10750-011-0935-6>.

2813 Schouten, S., Hopmans, E. C., Sinninghe Damsté, J. S., 2013. The organic geochemistry of glycerol dialkyl
2814 glycerol tetraether lipids: A review. *Org. Geochem.* 54, 19-61.

2815 Schreiber, K., Hauffe, T., Albrecht, C., Wilke, T., 2012. The role of barriers and gradients in differentiation
2816 processes of pyrgulinid microgastropods of Lake Ohrid. *Hydrobiologia* 682, 61-73.
2817 <http://doi.org/10.1007/s10750-011-0864-4>.

2818 Schüler, L., Hemp, A., Behling, H., 2014. Relationship between vegetation and modern pollen-rain along an
2819 elevational gradient on Kilimanjaro, Tanzania. *Holocene* 24, 702-713.

2820 Schultheiß, R., Van Bocxlaer, B., Wilke, T., Albrecht, C., 2009. Old fossils–young species: evolutionary history of
2821 an endemic gastropod assemblage in Lake Malawi, *Proc. R. Soc. Lond. B* 276, 2837-2846.
2822 <http://dx.doi.org/10.1098/rspb.2009.0467>.

2823 Schultheiß, R., Wilke, T., Jørgensen, A., Albrecht, C., 2011. The birth of an endemic species flock: demographic
2824 history of the *Bellamya* group (Gastropoda, Viviparidae) in Lake Malawi. *Biol. J. Linn. Soc.* 102, 130-143.
2825 <http://dx.doi.org/10.1111/j.1095-8312.2010.01574.x>.

2826 Seddon, A.W.R., Mackay, A.W., Baker, A.G. et al., 2014. Looking forward through the past: identification of 50
2827 priority research questions in palaeoecology. *J. Ecol.*, 102, 256-267. [http://dx.doi.org/10.1111/1365-](http://dx.doi.org/10.1111/1365-2745.12195)
2828 [2745.12195](http://dx.doi.org/10.1111/1365-2745.12195)).

2829 Shanahan, T. M., Peck, J. A., McKay, N., Heil Jr, C. W., King, J., Forman, S. L., Hoffmann, D. L., Richards, D. A.,
2830 Overpeck, J. T., Scholz, C., 2013. Age models for long lacustrine sediment records using multiple dating
2831 approaches – An example from Lake Bosumtwi, Ghana. *Quat. Geochronol.* 15, 47-60.
2832 <http://dx.doi.org/10.1016/j.quageo.2012.12.001>.

2833 Sherbakov, D. Y., 1999. Molecular phylogenetic studies on the origin of biodiversity in Lake Baikal. *Trends Ecol.*
2834 *Evol.* 14, 92-95. [http://dx.doi.org/10.1016/S0169-5347\(98\)01543-2](http://dx.doi.org/10.1016/S0169-5347(98)01543-2).

2835 Shi, J. J., Rabosky, D. L., 2015. Speciation dynamics during the global radiation of extant bats. *Evolution* 6, 1528-
2836 1545. <http://dx.doi.org/10.1111/evo.12681>.

2837 Silvestro, D., Schnitzler, J., Liow, L. H., Antonelli, A., Salamin, N., 2014. Bayesian estimation of speciation and
2838 extinction from incomplete fossil occurrence data. *Syst. Biol.* 63, 349-367.
2839 <http://dx.doi.org/10.1093/sysbio/syu006>.

2840 Smith, D. M., 2012. Exceptional preservation of insects in lacustrine environments. *Palaios* 27, 346-353.

2841 Slater, G. J., Harmon, L. J., Alfaro, M. E., 2012. Integrating fossils with molecular phylogenies improves inference
2842 of trait evolution. *Evolution* 66, 3931-3944. <http://dx.doi.org/10.1111/j.1558-5646.2012.01723.x>

2843 Snyder, J. A., Cherepanova, M. V., Bryan, A., 2013. Dynamic diatom response to changing climate 0-1.2Ma at
2844 Lake El'gygytgyn, Far East Russian Arctic. *Clim. Past* 9, 1309-1319.

2845 Sorenson, L., Santini, F., Alfaro, M. E., 2014. The effect of habitat on modern shark diversification. *J. Evol. Biol.*
2846 27, 1536-1548.

2847 Stadler, T., 2011a. Mammalian phylogeny reveals recent diversification rate shifts. *Proc. Natl. Acad. Sci. USA* 108,
2848 6187-6192. <http://dx.doi.org/10.1073/pnas.1016876108>.

2849 Stadler, T., 2011b. Inferring speciation and extinction processes from extant species data. *Proc. Natl. Acad. Sci.*
2850 *USA* 108, 16145-16146. <http://dx.doi.org/10.1073/pnas.1113242108>.

2851 Stager, J. C., Cocquyt, C., Bonnefille, R., Weyhenmeyer, C., Bowerman, N., 2009. A late Holocene paleoclimatic
2852 history of Lake Tanganyika, East Africa. *Quat. Res.* 72, 47-56. <http://dx.doi.org/10.1016/j.yqres.2009.04.003>.

2853 Stelbrink, B., Shirokaya, A. A., Clewing, C., Sitnikova, T. Y., Prozorova, L. A., Albrecht, C., 2015. Conquest of the
2854 deep, old and cold: an exceptional limpet radiation in Lake Baikal. *Biol. Lett.* 11, 20150321.
2855 <http://dx.doi.org/10.1098/rsbl.2015.0321>.

- 2856 Stevens, L., Djamali, M., Andrieu-Ponel, V., de Beaulieu, J.-L., 2012. Hydroclimatic variations over the last two
2857 glacial/interglacial cycles at Lake Urmī. Iran. J. Paleolimnol. 47, 645-660.
- 2858 Stockhecke, M., Kwiecien, O., Vigliotti, L., Anselmetti, F. S., Beer, J., Çağatay, M. N., Channell, J. E. T., Kipfer,
2859 R., Lachner, J., Litt, T., Pickarski, N., Sturm, M., 2014. Chronostratigraphy of the 600,000 year old continental
2860 record of Lake Van (Turkey), Quat. Sci. Rev. 104, 8-17. <http://dx.doi.org/10.1016/j.quascirev.2014.04.008>.
- 2861 Stoermer, E. F., Edlund, M. B., 1999. No paradox in the plankton? Diatom communities in large lakes. In:
2862 Mayama, S., Idei, M., Koizumi, I. (Eds.), Proceedings of the 14th International Diatom Symposium. Koeltz,
2863 Königstein, pp. 49-61.
- 2864 Storvall, H., Ramsköld, D., Sandberg, R., 2013. Efficient and comprehensive representation of uniqueness for next-
2865 generation sequencing by minimum unique length analyses. PLoS ONE 8, e53822.
2866 <http://dx.doi.org/10.1371/journal.pone.0053822>.
- 2867 Sulpizio, R., Alçiçek, M. C., Zanchetta, G., Solari, L., 2013. Recognition of the Minoan tephra in the Acigöl Basin,
2868 western Turkey: implications for inter-archive correlations and fine ash dispersal. J. Quat. Sci. 28, 329-335.
2869 <http://dx.doi.org/10.1002/jqs.2630>.
- 2870 Sulpizio, R., Zanchetta, G., D’Orazio, M., Vogel, H., Wagner, B., 2010. Tephrostratigraphy and tephrochronology
2871 of lakes Ohrid and Prespa, Balkans. Biogeosciences 7, 3273-3288. <http://dx.doi.org/10.5194/bg-7-3273-2010>.
- 2872 Sutton, P. A., Rowland, S. J., 2012. High temperature gas chromatography–time-of-flight-mass spectrometry
2873 (HTGC–ToF-MS) for high-boiling compounds. J. Chromatogr. A 1243, 69-80.
- 2874 Sykes, G. A., Collins, M. J., Walton, D. I., 1995. The significance of a geochemically isolated intracrystalline
2875 fraction within biominerals. Org. Geochem. 23, 1059-1065.
- 2876 Taft, L., Wiechert, U., Riedel, F., Weynell, M., Zhang, H., 2012. Sub-seasonal oxygen and carbon isotope
2877 variations in shells of modern *Radix* sp. (Gastropoda) from the Tibetan Plateau: Potential of a new archive for
2878 palaeoclimatic studies. Quaternary Sci. Rev. 34, 44-56. <http://dx.doi.org/10.1016/j.quascirev.2011.12.006>.
- 2879 Takahata, N., 2007. Molecular Clock: An Anti-neo-Darwinian Legacy. Genetics 176, 1-6.
- 2880 Takahashi, T., Moreno, E., 2015. A RAD-based phylogenetics for *Orestias* fishes from Lake Titicaca. Mol.
2881 Phylogenet. Evol. 93, 307-317.
- 2882 Takemura, K., Haraguchi, T., Kusumoto, S., Itoh, Y., 2013. Tectonic basin formation in and around Lake Biwa,
2883 central Japan. In: Ito, Y. (Ed.), Mechanism of sedimentary basin formation - multidisciplinary approach on
2884 active plate margins. InTech, Rijeka, pp. 209-229.
- 2885 Talbot, H. M., Watson, D. F., Pearson, E. J., Farrimond, P., 2003. Diverse biohopanoid compositions of non-marine

- 2886 sediments. *Org. Geochem.* 34, 1353-1371.
- 2887 Talbot, M. R., Kelts, K., 1986. Primary and diagenetic carbonates in the anoxic sediments of Lake Bosumtwi,
2888 Ghana. *Geology* 14, 912-916.
- 2889 Tauxe, L., 1993. Sedimentary records of relative paleointensity of the geomagnetic field: theory and practice. *Rev.*
2890 *Geophys.* 31, 319-319.
- 2891 Taviani, M., Beu, A. G., Jonkers, H. A., 2000. Macrofossils from CRP-2/2A, Victoria Land Basin, Antarctica.
2892 *Terra Antarctica* 7, 513-526.
- 2893 Teranes, J. L., McKenzie, J. A., 2001. Lacustrine oxygen isotope record of 20th-century climate change in central
2894 Europe: evaluation of climatic controls on oxygen isotopes in precipitation. *J. Paleolimnol.* 26, 131-146.
- 2895 Thomas, C., Ionescu, D., Ariztegui, D., DSDDP Scientific Team, 2014. Archaeal populations in two distinct
2896 sedimentary facies of the subsurface of the Dead Sea. *Mar. Genomics* 17, 53-62.
- 2897 Thomas, C., Ionescu, D., Ariztegui, D., DSDDP Scientific Team, 2015. Climate ruling life in a hypersaline
2898 subsurface: Identifying microbial populations in the Dead Sea seep sediments by 16S rRNA gene sequence
2899 analysis. *Geobiology* 13, 546-561. <http://dx.doi.org/10.1111/gbi.12151>.
- 2900 Thomas, G. H., Freckleton, R. P., 2012. MOTMOT: models of trait macroevolution on trees. *Methods Ecol. Evol.*
2901 3, 145-151. <http://dx.doi.org/10.1111/j.2041-210X.2011.00132.x>.
- 2902 Tierney, J. E., Russell, J. M., Huang, Y., 2010. A molecular perspective on Late Quaternary climate and vegetation
2903 change in the Lake Tanganyika basin, East Africa. *Quat. Sci. Rev.* 29, 787-800.
- 2904 Tierney, J. E., Schouten, S., Pitcher, A., Hopmans, E. C., Sinninghe Damsté, J. S., 2012. Core and intact polar
2905 glycerol dialkyl glycerol tetraethers (GDGTs) in Sand Pond, Warwick, Rhode Island (USA): insights into the
2906 origin of lacustrine GDGTs. *Geochim. Cosmochim. Acta* 77, 561-581.
- 2907 Tolu, J., Gerber, L., Boily, J. F., Bindler, R., 2015. High-throughput characterization of sediment organic matter by
2908 pyrolysis-gas chromatography/mass spectrometry and multivariate curve resolution: a promising analytical tool
2909 in (paleo)limnology. *Anal. Chim. Acta* 880, 93-102.
- 2910 Tomlinson, E. L., Smith, V. C., Albert, P. G., Aydar, E., Civetta, L., Cioni, R., Çubukçu, E., Gertisser, R., Isaia, R.,
2911 Menzies, M. A., Orsi, G., Rosi, M., Zanchetta, G., 2015. The major and trace element glass compositions of the
2912 productive Mediterranean volcanic sources: tools for correlating distal tephra layers in and around Europe.
2913 *Quat. Sci. Rev.* 118, 48-66. <http://dx.doi.org/10.1016/j.quascirev.2014.10.028>.
- 2914 Towe, K. M., 1980. Preserved organic ultrastructure: an unreliable indicator for Paleozoic amino acid
2915 biogeochemistry. In: Hare, P. E., Hoering, T. C., King, K., Jr. (Eds.), *Biogeochemistry of Amino Acids*. Wiley,

2916 New York, pp. 65-74.

2917 Trajanovski, S., Albrecht, C., Schreiber, K., Schultheiß, R., Stadler, T., Benke, M., Wilke, T., 2010. Testing the
2918 spatial and temporal framework of speciation in an ancient lake species flock: the leech genus *Dina* (Hirudinea:
2919 Erpobdellidae) in Lake Ohrid. *Biogeosciences* 7, 3387-3402. <http://dx.doi.org/10.5194/bg-7-3387-2010>.

2920 Trondman, A.-K., Gaillard, M.-J., Mazier, F., Sugita, S., Fyfe, R., Nielsen, A. B., Twiddle, C., Barratt, P., Birks, H.
2921 J. B., Bjune, A. E., Björkman, L., Broström, A., Caseldine, C., David, R., Dodson, J., Dörfler, W., Fischer, E.,
2922 Van Geel, B., Giesecke, T., Hultberg, T., Kalnina, L., Kangur, M., Van der Knaap, P., Koff, T., Kuneš, P.,
2923 Lagerås, P., Latałowa, M., Lechterbeck, J., Leroyer, C., Leydet, M., Lindbladh, M., Marquer, L., Mitchell, F. J.
2924 G., Odgaard, B. V., Peglar, S. M., Persson, T., Poska, A., Rösch, M., Seppä, H., Veski, S., Wick, L., 2015.
2925 Pollen-based quantitative reconstructions of Holocene regional vegetation cover (plant-functional types and
2926 land-cover types) in Europe suitable for climate modelling. *Glob. Chang. Biol.* 21, 676-697.

2927 Tzedakis, P. C., Andrieu, V., Birks, H. J. B., De Beaulieu, J.-L., Crowhurst, S., Follieri, M., Hooghiemstra, H.,
2928 Magri, D., Reille, M., Sadori, L., Shackleton, N. J., Wijmstra, T. A., 2001. Establishing a terrestrial
2929 chronological framework as a basis for biostratigraphical comparisons. *Quat. Sci. Rev.* 20, 1583-1592.

2930 Tzedakis, P. C., Andrieu, V., De Beaulieu, J.-L., Crowhurst, S., Follieri, M., Hooghiemstra, Magri, D., Reille, M.,
2931 Sadori, L., Shackleton, N. J., Wijmstra, T. A., 1997. Comparison of terrestrial and marine records of changing
2932 climate of the last 500,000 years. *Earth Planet. Sci. Lett.* 150, 171-176.

2933 Urey, H. C., Lowenstam, H. A., Epstein, S., McKinney, C. R., 1951. Measurement of palaeotemperatures and
2934 temperatures of the Upper Cretaceous of England, Denmark and Southeastern United States. *Geol. Soc. Am.*
2935 *Bull.* 62, 399-416.

2936 Van Bocxlaer, B., Clewing, C., Etimosundja, J.-P. M., Kankonda, A., Ndeo, O. W., Albrecht, C., 2015. Recurrent
2937 camouflaged invasions and dispersal of an Asian freshwater gastropod in tropical Africa. *BMC Evol. Biol.* 15,
2938 1-18. <http://dx.doi.org/10.1186/s12862-015-0296-2>.

2939 Van Bocxlaer, B., Hunt, G., 2013. Morphological stasis in an ongoing gastropod radiation from Lake Malawi. *Proc.*
2940 *Natl. Acad. Sci. USA* 110, 13892-13897. <http://dx.doi.org/10.1073/pnas.1308588110>.

2941 Van Bocxlaer, B., Van Damme, D., Feibel, C., 2008. Gradual versus punctuated equilibrium evolution in the
2942 Turkana Basin molluscs: Evolutionary events or biological invasions. *Evolution* 62, 511-520. [http://dx.doi.org/](http://dx.doi.org/10.1111/j.1558-5646.2007.00296.x)
2943 [10.1111/j.1558-5646.2007.00296.x](http://dx.doi.org/10.1111/j.1558-5646.2007.00296.x).

2944 Vogel, H., Rosén, P., Wagner, B., Melles, M., Persson, P., 2008. Fourier transform infrared spectroscopy, a new
2945 cost-effective tool for quantitative analysis of biogeochemical properties in long sediment records. *J.*

2946 Paleolimnol. 40, 689-702.

2947 Vogel, H., Wagner, B., Zanchetta, G., Sulpizio, R., Rosén, P., 2010b. A paleoclimate record with
 2948 tephrochronological age control for the last glacial-interglacial cycle from Lake Ohrid, Albania and Macedonia.
 2949 J. Paleolimnol. 41, 407-430.

2950 Vogel, H., Zanchetta, G., Sulpizio, R., Wagner, B., Nowaczyk, N., 2010a. A tephrostratigraphic record for the last
 2951 glacial-interglacial cycle from Lake Ohrid, Albania and Macedonia. J. Quat. Sci. 25, 320-338.
 2952 <http://dx.doi.org/10.1002/jqs.1311>.

2953 Volkman, J. K., Barrett, S. M., Blackburn, S. I., Mansour, M. P., Sikes, E. L., Gelin, F., 1998. Microalgal
 2954 biomarkers: a review of recent research developments. Org. Geochem. 29, 1163-1179.

2955 von Rintelen, T., Wilson, A. B., Meyer, A., Glaubrecht, M., 2004. Escalation and trophic specialization drive
 2956 adaptive radiation of freshwater gastropods in ancient lakes on Sulawesi, Indonesia. Proc. R. Soc. Lond. B Biol.
 2957 Sci. 271, 2541-2549. <http://dx.doi.org/10.1098/rspb.2004.2842>.

2958 von Rintelen, K., Glaubrecht, M., Schubart, C. D., Wessel, A., von Rintelen, T., 2010. Adaptive radiation and
 2959 ecological diversification of Sulawesi's ancient lake shrimps. Evolution 64, 3287-3299.

2960 Vuillemin, A., Ariztegui, D., De Coninck, A., Lücke, A., Mayr, C., Schubert, C., PASADO Scientific Team, 2013b.
 2961 Origin and significance of diagenetic concretions in sediments of Laguna Potrok Aike, southern Argentina. J.
 2962 Paleolimnol. 50, 275-291.

2963 Vuillemin, A., Ariztegui, D., Lücke, A., Mayr, C., PASADO Scientific Team, 2014a. Paleoenvironmental
 2964 conditions define current sustainability of microbial populations in Laguna Potrok Aike sediments, Argentina.
 2965 Aquat. Sci. 76, 101-114.

2966 Vuillemin, A., Ariztegui, D., Nobbe, G., Schubert, C., PASADO Scientific Team, 2014b. Influence of
 2967 methanogenic populations in Holocene lacustrine sediments revealed by clone libraries and fatty acid
 2968 biogeochemistry. Geomicrobiol. J. 31, 285-298.

2969 Vuillemin, A., Ariztegui, D., PASADO Scientific Team, 2013a. Geomicrobiological investigations in subsaline
 2970 maar lake sediments over the last 1500 years. Quat. Sci. Rev. 71, 119-130.

2971 Vuillemin, A., Ariztegui, D., Vasconcelos, C., PASADO Scientific Drilling Party, 2010. Establishing sampling
 2972 procedures in lake cores for subsurface biosphere studies: Assessing in situ microbial activity. Sci. Drill. 10, 35-
 2973 39.

2974 Wagner, B., Leng, M., Wilke, T., Böhm, A., Panagiotopoulos, K., Vogel, H., Lacey, J., Zanchetta, G., Sulpizio, R.,
 2975 2014b. Distinct lake level lowstand in Lake Prespa (SE Europe) at the timing of the 74 (75) ka Toba eruption.

- 2976 Clim. Past 10, 261-267. <http://dx.doi.org/10.5194/cp-10-261-2014>.
- 2977 Wagner, B., Lotter, A. F., Nowaczyk, N., Reed, J.M., Schwalb, A., Sulpizio, R., Valsecchi, V., Wessels, M.,
2978 Zanchetta, G., 2009. A 40,000-year record of environmental change from ancient Lake Ohrid (Albania and
2979 Macedonia). *J. Paleolimnol.* 41, 407-430.
- 2980 Wagner, B., Wilke, T., Krastel, S., Zanchetta, G., Sulpizio, R., Reicherter, K., Leng, M., Grazhdani, A.,
2981 Trajanovski, S., Francke, A., Lindhorst, K., Levkov, Z., Cvetkoska, A., Reed, J., Zhang, X., Lacey, J., Wonik,
2982 T., Baumgarten H., 2014a. The SCOPSCO drilling project recovers more than 1.2 million years of history from
2983 Lake Ohrid. *Sci. Drill.* 17, 19-29. <http://dx.doi.org/10.5194/sd-17-19-2014>.
- 2984 Wagner, C. E., Harmon, L. J., Seehausen, O., 2014c. Cichlid species-area relationships are shaped by adaptive
2985 radiations that scale with area. *Ecol. Lett.* 17, 583-592. <http://dx.doi.org/10.1111/ele.12260>.
- 2986 Wagner, C. E., Harmon, L. J., Seehausen, O., 2012. Ecological opportunity and sexual selection together predict
2987 adaptive radiation. *Nature* 487, 366-369. <http://dx.doi.org/10.1038/nature11144>.
- 2988 Warnock, R. C., Parham, J. F., Joyce, W. G., Lyson, T. R., Donoghue, P. C., 2015. Calibration uncertainty in
2989 molecular dating analyses: there is no substitute for the prior evaluation of time priors. *Proc. R. Soc. Lond. B*
2990 282, 20141013.
- 2991 Weber, M. E., Niessen, F., Kuhn, G., Wiedicke, M., 1997. Calibration and application of marine sedimentary
2992 physical properties using a multi-sensor core logger. *Mar. Geol.* 136, 151-172.
- 2993 Wehmiller, J. F., 2012. United States Quaternary coastal sequences and molluscan racemization geochronology –
2994 What have they meant for each other over the past 45 years? *Quat. Geochronol.* 16, 3-20.
- 2995 Weijers, J. W. H., Schouten, S., van den Donker, J. C., Hopmans, E. C., Sinninghe Damsté, J. S., 2007.
2996 Environmental controls on bacterial tetraether membrane lipid distribution in soils. *Geochim. Cosmochim. Acta*
2997 71, 703-713.
- 2998 Weiss, J. D., Cotterill, F. P. D., Schlieven, U. K., 2015. Lake Tanganyika – a “melting pot” of ancient and young
2999 cichlid lineages (Teleostei: Cichlidae)? *PLoS ONE* 10, e0125043.
3000 <http://dx.doi.org/10.1371/journal.pone.0125043>.
- 3001 Welander, P. V., Coleman, M. L., Sessions, A. L., Summons, R. E., Newman, D. K., 2012. Identification of a
3002 methylase required for 2-methylhopanoid production and implications for the interpretation of sedimentary
3003 hopanes. *Proc. Natl. Acad. Sci. USA* 107, 8537-8542.
- 3004 Wesselingh, F. P., 2007. Long-lived lake molluscs as island faunas: a bivalve perspective. In: Renema, W. (Ed.),
3005 Biogeography, time and place: distributions, barriers and islands. Springer, Dordrecht, pp. 275-314.

3006 Whitman, W. B., Coleman, D. C., Wiebe, W. J., 1998. Prokaryotes: The unseen majority. *Proc. Natl. Acad. Sci.*
3007 *USA* 95, 6578-6583.

3008 Wick, L., Lemcke, G., Sturm, M., 2003. Evidence of Lateglacial and Holocene climatic change and human impact
3009 in eastern Anatolia: high-resolution pollen, charcoal, isotopic and geochemical records from the laminated
3010 sediments of Lake Van, Turkey. *Holocene* 13, 665-675.

3011 Wilke, T., 2004. How dependable is a non-local molecular clock? A reply to Hausdorf et al., 2003. *Mol.*
3012 *Phylogenet. Evol.* 30, 835-840.

3013 Wilke, T., Schultheiß, R., Albrecht, C., 2009. As time goes by: a simple fool's guide to molecular clock approaches
3014 in invertebrates. *Am. Malacol. Bull.* 27, 25-45. <http://dx.doi.org/10.4003/006.027.0203>.

3015 Wilkie, K. M. K., Chaplignin, B., Meyer, H., Burns, S., Petsch, S., Brigham-Grette, J., 2013. Modern isotope
3016 hydrology and controls on δD of plant leaf waxes at Lake El'gygytgyn, NE Russia. *Clim. Past* 9, 335-352.

3017 Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., Lorenzen, E. D., Vestergård, M.,
3018 Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L. S., Pearman, P. B., Cheddadi, R.,
3019 Murray, D., Bråthen, K. A., Yoccoz, N., Binney, H., Cruaud, C., Wincker, P., Goslar, T., Alsos, I. G.,
3020 Bellemain, E., Brysting, A. K., Elven, R., Sønstebø, J. H., Murton, J., Sher, A., Rasmussen, M., Rønn, R.,
3021 Mourier, T., Cooper, A., Austin, J., Möller, P., Froese, D., Zazula, G., Pompanon, F., Rioux, D., Niderkorn, V.,
3022 Tikhonov, A., Savvinov, G., Roberts, R. G., MacPhee, R. D., Gilbert, M. T., Kjær, K. H., Orlando, L.,
3023 Brochmann, C., Taberlet, P., 2014. Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature* 506,
3024 47-51. <http://dx.doi.org/10.1038/nature12921>.

3025 Willerslev, E., Cappellini, E., Boomsma, W., Nielsen, R., Hebsgaard, M. B., Brand, T. B., Hofreiter, M., Bunce,
3026 M., Poinar, H. N., Dahl-Jensen, D., Johnsen, S., Steffensen, J. P., Bennike, O., Schwenninger, J.-L., Nathan, R.,
3027 Armitage, S., De Hoog, C.-J., Alfimov, V., Christl, M., Beer, J., Muscheler, R., Barker, J., Sharp, M., Penkman,
3028 K. E. H., Haile, J., Taberlet, P., Gilbert, M. T. P., Casoli, A., Campani, E., Collins, M. J., 2007. Ancient
3029 Biomolecules from Deep Ice Cores Reveal a Forested Southern Greenland. *Science* 317, 111-114.

3030 Williamson, P. G., 1981. Paleontological documentation of speciation in Cenozoic mollusks from Turkana Basin.
3031 *Nature* 293, 437-443.

3032 Williamson, P. G., 1985. Evidence for an early Plio-Pleistocene rainforest expansion in East Africa. *Nature* 315,
3033 487-489.

3034 Wilson, A. B., Glaubrecht, M., Meyer, A., 2004. Ancient lakes as evolutionary reservoirs: evidence from the
3035 thalassoid gastropods of Lake Tanganyika. *Proc. R. Soc. Lond. B* 271, 529-536.

- 3036 Winder, M., Reuter, J. E., Schladow, S. G., 2009. Lake warming favours small-sized planktonic diatom species.
3037 Proc. R. Soc. Lond. B 276, 427-435. <http://dx.doi.org/0.1098/rspb.2008.1200>.
- 3038 Wulf, S., Kraml, M., Keller, J., 2008. Towards a detailed distal tephrostratigraphy in the Central Mediterranean:
3039 The last 20,000 yrs record of Lago Grande di Monticchio. *J. Volcanol. Geotherm. Res.* 177, 118-132.
3040 <http://dx.doi.org/10.1016/j.jvolgeores.2007.10.009>.
- 3041 Wysocka, A., Grabowski, M., Sworobowicz, L., Mamos, T., Burzyński, A., Sell, J., 2014. Origin of the Lake Ohrid
3042 gammarid species flock: ancient local phylogenetic lineage diversification. *J. Biogeogr.* 41, 1758-1768.
3043 <http://dx.doi.org/10.1111/jbi.12335>.
- 3044 Young, K. A., Snoeks, J., Seehausen, O., 2009. Morphological diversity and the roles of contingency, chance and
3045 determinism in african cichlid radiations. *PloS ONE* 4, e4740. <http://dx.doi.org/10.1371/journal.pone.0004740>.
- 3046 Zanchetta, G., Regattieri, E., Giaccio, B., Wagner, B., Sulpizio, R., Francke, A., Vogel, H., Sadori, L., Masi, A.,
3047 Sinopoli, G., Lacey, J.H., Leng, M. L., Leicher, N., 2015. Aligning MIS5 proxy records from Lake Ohrid
3048 (FYROM) with independently dated Mediterranean archives: implications for core chronology. *Biogeosci.*
3049 *Discuss.* 12, 15461–15493. <http://dx.doi.org/10.5194/bgd-12-15461-2015>.
- 3050 Zarzycki, P. K., Portka, J. K., 2015. Recent advances in hopanoids analysis: Quantification protocols overview,
3051 main research targets and selected problems of complex data exploration. *J. Steroid Biochem. Mol. Biol.* 153, 3-
3052 26.
- 3053 Zhang, X. S., Reed, J. M., Lacey, J. H., Francke, A., Leng, M. J., Levkov, Z., Wagner, B., 2016. Complexity of
3054 diatom response to Lateglacial and Holocene climate and environmental change in ancient, deep and
3055 oligotrophic Lake Ohrid (Macedonia and Albania). *Biogeosciences* 13, 1351-1365.
3056 <http://dx.doi.org/10.5194/bg-13-1351-2016>.
- 3057 Zuckerkandl, E., Pauling, L., 1965. Evolutionary divergence and convergence in proteins. In: Bryson, V., Vogel H.
3058 J. (Eds.), *Evolving Genes and Proteins*. Academic Press, New York, pp. 97-166.