CORE

Metadata, citation and similar papers at core.ac.uk

Provided by Northumbria Research Link

Northumbria Research Link

Citation: Bremond, Laurent, Favier, Charly, Ficetola, Gentile Francesco, Tossou, Monique G., Akouégninou, Akpovi, Gielly, Ludovic, Giguet-Covex, Charline, Oslisly, Richard and Salzmann, Ulrich (2017) Five thousand years of tropical lake sediment DNA records from Benin. Quaternary Science Reviews, 170. pp. 203-211. ISSN 0277-3791

Published by: Elsevier

URL: https://doi.org/10.1016/j.quascirev.2017.06.025 <https://doi.org/10.1016/j.quascirev.2017.06.025>

This version was downloaded from Northumbria Research Link: http://nrl.northumbria.ac.uk/31208/

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: http://nrl.northumbria.ac.uk/policies.html

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

www.northumbria.ac.uk/nrl

Five thousand years of tropical lake sediment DNA records from Benin

Bremond L.[1,2], Favier C.[1], Ficetola G.F.[3,4,5], Tossou M.G.[6], Akouégninou A.[6], Gielly L.[3,4], Giguet-Covex C.[7], , Oslisly. R.[8], Salzmann U.[9]

1 Institut des Sciences de l'Évolution - Montpellier, UMR 5554 CNRS-IRD-Université Montpellier-EPHE, Montpellier, France.

2 École Pratique des Hautes Études, PSL Research University, 4-14 rue Ferrus, 75014 Paris, France.

3 University Grenoble Alpes, LECA, F-38000 Grenoble, France

4 [4]CNRS, LECA, F-38000 Grenoble, France.

5 Department of Biosciences, Università degli Studi di Milano, Via Celoria 26, 20133, Milan, Italy.

6 National Herbarium, Department of Botany and Plant Biology, Faculty of Sciences and Technology (FAST) University of Abomey-Calavi (UAC) Cotonou, Benin.

7 EDYTEM, Université de Savoie, CNRS, Pôle Montagne, 73376 Le Bourget du Lac, France.

8 UMR PALOC Institut de Recherche pour le Développement et Agence Nationale des Parcs Nationaux, BP 20379 Libreville, Gabon.

9. Department of Geography, Faculty of Engineering and Environment, Northumbria University, Ellison Building, Newcastle upon Tyne, NE1 8ST, UK

Corresponding author: Laurent Bremond

25 E-mail address: laurent.bremond@umontpellier.fr
26 Postal address: Institut des Sciences de l'Evolution - Montpellier
27 UMR5554 UM – CNRS – IRD - EPHE
28 Université de Montpellier
29 CC 061 34095 MONTPELLIER Cedex 05
30
31 Contributors: LB and CF designed the study; RO, MGT and AA contributed to the design
32 and permit technical support; GFF, LG, LB performed sampling and DNA extraction; GFF
33 LG, CF analyzed the data; LB, GFF, CF, US, CGC performed data interpretation; LB drafted
34 the article, and all authors critically revised the article and approved the version to be
35 published.
36

37 **Abstract**

38 Until now, sedimentary DNA (*sed*DNA) studies have only focused on cold and temperate
39 regions were DNA is relatively well preserved. Consequently, the tropics, where
40 vegetation is hyperdiverse and natural archives are rare, have been neglected and
41 deserve attention. In this study, we used next-generation sequencing to barcode *sed*DNA
42 from Lake Sele, localized in the tropical lowlands of Benin (Africa), and compared the
43 taxonomic diversity detected by DNA analyses with pollen assemblages. Plant *sed*DNA
44 was successfully amplified from 33 of the 34 successfully extracted samples. In total, 43
45 taxa were identified along the 5,000 years spanned by the sediment: 22 taxa were
46 identified at the family level and 21 at the genus level. The plant diversity recovered
47 through *sed*DNA from Lake Sele showed a specific local signal and limited overlapping
48 with pollen. Introduced plants, grown and cultivated close to the water, such as sweet
49 potato, were also well recorded by *sed*DNA. It appears, therefore, to be a promising

50 approach to studying past diversity in tropical regions, and could help in tracking the
51 introduction and history of agriculture. This is the first time this method has been used
52 in the field of domestication and dissemination of several specific crops, and the results
53 are very encouraging.

54

55 **Keywords:** *sed*DNA, metabarcoding, tropical, Africa, vegetation dynamics, plant
56 diversity, pollen

57

58 **Introduction**

59 Plant sedimentary DNA (*sed*DNA) has recently emerged as a promising proxy for
60 paleoecological reconstructions over thousands of years, complementary to pollen and
61 macrofossils analysis (Thomsen and Willerslev, 2015; Parducci et al., 2017). However,
62 until now, paleoenvironmental plant *sed*DNA analyses have essentially been conducted
63 in temperate and boreal zones (Willerslev et al., 2003; Haile et al., 2009; Jørgensen et al.,
64 2012; Parducci et al., 2013; Giguet-Covex et al., 2014; Parducci et al., 2017). The oldest
65 *sed*DNA records were recovered from the permafrost of Siberia (Willerslev et al., 2003)
66 and Northern Scandinavia (Parducci et al., 2012). In more temperate zones, the
67 published results generally span over the last 10,000 years (Anderson-Carpenter et al.,
68 2011; Giguet-Covex et al., 2014; Pansu et al., 2015; Alsos et al., 2016), as the studied
69 lakes have accumulated sediments since the last deglaciation. While the taxonomic
70 resolution of *sed*DNA is not necessarily better than pollen or macroremain counts
71 (Matisoo-Smith et al., 2008; Parducci et al., 2013; Pedersen et al., 2013), they
72 nevertheless allow the detection of specific taxa either poorly or not recorded by
73 classical paleoecological methods. The analysis of plant *sed*DNA has enabled multiple
74 improvements to our paleoecological knowledge, such as the discovery that coniferous
75 trees survived the last glaciation in Northern Scandinavia despite this not being detected
76 by pollen analysis (Parducci et al., 2012), and the description of long-term vegetation
77 changes determined by human activities and pasture (Giguet-Covex et al., 2014; Pansu
78 et al., 2015).

79 Few studies have explored *sed*DNA in tropical regions (Epp et al., 2010; Stoof-
80 Leichsenring et al., 2012), and only one, from a high-altitude lake, analyzed long-term
81 variations in terrestrial plant DNA (Boessenkool et al., 2014). The initial focus on high-
82 latitude or high-altitude areas is related to the fact that DNA degradation increases at

83 high temperatures (Allentoft et al., 2012), so the absence of publications on plant
84 *sed*DNA analyses in low-altitude and latitude sites might merely be due to the absence of
85 trials or a failure of these trials (Parducci et al., 2017). Despite the relatively short time
86 period examined, the recovery of diatom *sed*DNA dating back 200 years by Stoof-
87 Leichsenring *et al.* (2011) in Lake Naivasha, Kenya, is particularly interesting as water
88 temperatures rarely fall below 20° C (in Britton et al., 2007). In the two high-altitude
89 lakes studied by Boessenkool *et al.* (2014) *sed*DNA sequences were obtain from 4,500-
90 year-old samples. With the help of a reference library covering the majority of the afro-
91 alpine flora, the authors were able to identify several taxa at a better/more precise
92 taxonomic level than pollen analyses, but with fewer species detected. If DNA can be
93 preserved in sediments from warm climates, the method could be an important proxy to
94 identify early agriculture and farming in the tropics. In Africa, for example, the onset of
95 the cultivation of pearl millet (*Pennisetum glaucum*) or sorghum (*Sorghum bicolor*) is
96 poorly understood (Fuller et al., 2014) because traditional bioproxies such as pollen,
97 and phytoliths are inadequate (phytolith and pollen of African grasses cannot be
98 differentiated into cereals and wild grasses), and charred grains are rare. *Sed*DNA may
99 also be a valuable tool to complement and alleviate the time constraints of
100 phylogeography studies, the study of the geographic distribution of genetic lineages
101 within species (e.g. for Aucoumea klaineana Born et al., 2008), or to track the history of
102 specific vegetation types badly recovered by more conventional methods, for instance
103 the Marantaceae forests (Cuni-Sanchez et al., 2016).
104 The objective of this study is to explore the preservation of plant DNA in the sediments
105 of a low-elevation tropical lake over the last 5,000 years, and to test the pertinence of
106 *sed*DNA analysis to record past changes in plant diversity and agriculture practices on
107 the shoreline of the lake. The DNA of wild and cultivated plants is generally incorporated

108 into the soils (Yoccoz et al., 2012), and may be transported to the lake by erosion during
109 rainfall. We chose Lake Sele (south of Benin) because the shorelines are flat, with large
110 parts being used for seasonal agriculture during the dry season when the lake level is
111 low. Furthermore, its sediments have already been studied for fossil pollen content
112 (Salzmann and Hoelzmann, 2005), allowing comparison between these records and the
113 *sed*DNA.
114

115 **Material and Method**

116 **Study site**

117 Lake Sele in Benin is located about 1 km east of the Ouémé River (7° 9'19.29"N,
118 2°26'25.57"E; Figure 1) at an altitude of less than 10 m above sea level, with a mean
119 annual temperature of 28° C (Hijmans et al., 2004); and though it is not directly fed by
120 the river it is within the same riverbed. The lake is elongated and shallow at around. 2.5
121 km in length, 1 km wide (at its maximum point), and with a maximum water depth of
122 <1.5m during the dry season (Salzmann and Hoelzmann, 2005). Nowadays, during the
123 dry season, the shorelines are partially covered with Cyperaceae and the water hyacinth
124 *Eichhornia crassipes* (Pontederiaceae; introduced from South America) (Salzmann and
125 Hoelzmann, 2005). The shorelines are generally cultivated when the lake level is low.
126 Agricultural parcels are present all around the lake forming a belt ~200-300m wide,
127 which are partially or totally submerged during the wet season. The main cultures are
128 sweet potatoes (*Ipomoea batatas*, Convolvulaceae), the red variety planted close to the
129 water and the white variety further away; also cultivated away from the shoreline are
130 cassava (*Manihot esculenta*, Euphorbiaceae), maize (*Zea mays*, Poaceae), and peanuts
131 (*Arachis hypogaea*, Fabaceae). Pigeon pea (*Cajanus cajan*, Fabaceae), bananas (*Musa sp.*,
132 Musaceae), tomatoes (*Solanum lycopersicum*, Solanaceae), and other legumes are

133 cultivated in the upper part of the shoreline, along with coconut (*Cocos nucifera*,
134 Arecaceae) and oil palms (*Elaeis guinensis*, Arecaceae). The natural vegetation belongs to
135 the Guinean transition zone with its mosaic of forests and savannas (White, 1983). The
136 main trees of the residual forest stands, namely *Triplochiton scleroxylon* (abachi,
137 Sterculiaceae), *Celtis* spp. (Cannabaceae), and Ulmaceae*, Milicia excelsa* (iroko,
138 Moraceae) are presented in Salzmann and Hoelzmann (2005).
139

140 **Sampling**

141 We collected five meters of sediment cores from the deepest part of the lake (N7.15537°
142 E2.44106°). The two upper meters of the sediment were collected with an Uwitec
143 Gravity corer (63 mm inner diameter), in order to get the longest possible section in one
144 tube to minimize contamination; the three lowers meters were taken with a modified
145 Livingston piston corer (47 mm inner diameter), which permitted us to reach the
146 deepest sediments in one meter sections. The whole sediment core was kept closed in
147 either plastic (Uwitec) or aluminum (Livingston) tubes and stored for several days in an
148 air-conditioned room in Benin, before being moved to a cold chamber (4° C) in
149 Montpellier, France. Sampling for DNA extraction was performed one month later after
150 arrival in France. Seven AMS ^{14}C analyses were carried out on terrestrial macroremains
151 (seeds and piece of plant leafs) by the Poznan Radiocarbon Laboratory. The ^{14}C ages
152 were converted to calendar years using CLAM software (Blaauw, 2010) and the age-
153 depth model was generated using BACON software (Blaauw and Christen, 2011). A
154 hiatus of sedimentation was detected between 4,400 years cal. BP and 3,100 years cal.
155 BP, as previously indicated by Salzmann and Hoelzmann (2005) (AMS ^{14}C dates and age-
156 depth model in supplementary data: Appendix S3).
157

158 **Samples for DNA analysis**

159 From the sediment core, we sampled 50 slices at 2 cm thick. Sampling of core slices was
160 carried out at the University of Savoie. To avoid DNA contamination, particularly with
161 regards to water circulation along the coring tube, 10 mm was removed from the edge of
162 each sediment slice. The tools (knives, spoons, pliers) were cleaned after each sampling
163 with distilled water, then alcohol and inflamed. DNA extraction, PCR set-up and PCR
164 amplification steps were later performed in three separate rooms at the University
165 Grenoble Alpes, which is specifically dedicated to ancient DNA studies. *Sed*DNA
166 extraction targeted extracellular DNA (Pansu et al., 2015). For each sediment slice, we
167 mixed approximately 15 g of sediment with 15 ml of saturated phosphate buffer
168 (Na2HPO4; 0.12 M; pH ≈ 8) for 15 minutes. The mixture was then centrifuged (10 min at
169 10000 g). The resulting supernatant (12 ml) was transferred to Amicon ® Ultra-15 10K
170 Centrifugal Filter Devices (Millipore) and centrifuged (20 min at 4000 g) to concentrate
171 *sed*DNA. Of the resulting concentrate, 400µl were kept as a starting material for the
172 following extraction steps, using the NucleoSpin® Soil kit (Macherey-Nagel, Düren,
173 Germany), skipping the cell lysis step and following the manufacturer's instructions
174 (Taberlet et al., 2012b). Three extraction controls were performed.
175 The *sed*DNA extractions and amplifications were performed in laboratories specifically
176 dedicated to ancient DNA studies in the University Grenoble-Alpes, France.
177

178 **DNA amplification and high-throughput sequencing**

179 Plant DNA was amplified with the universal primers "g" and "h" (Taberlet et al., 2007),
180 targeting the short and variable P6 loop region of the chloroplast trnL (UAA) intron.
181 These primers are highly conserved in angiosperms and gymnosperms, and are thus of
182 interest for plant metabarcoding (Taberlet et al., 2007). To assign sequence reads to the

183 relevant sample, 8 bp tags (with at least 5 differences between them) were added to the
184 5' end of primers (Binladen et al., 2007; Valentini et al., 2009). DNA amplifications were
185 carried out in a final volume of 30 µl containing 3 µl of diluted DNA extract. The
186 amplification mixture contained 1.2 U of AmpliTaq Gold® DNA Polymerase (Applied
187 Biosystems), 15 mM Tris-HCl, 50 mM KCl, 2 mM of MgCl2, 0.2 mM of each dNTP, 0.2 µM
188 of each primer, and 4.8 µg of bovine serum albumin (BSA, Roche Diagnostic). After 10
189 minutes at 95° C for polymerase activation, the PCR mixture underwent 45 cycles of 30 s
190 at 95 °C, 30 s at 50 °C, and 1 min at 72 °C, followed by a final elongation step (7 min at
191 72 °C). In order to reduce the rate of false negatives (Ficetola et al., 2015) PCR products
192 were then purified. All the samples were identified by unique 8 bp tags at the 5' end of
193 the primers, which allowed the identification of the products of each PCR. This allowed
194 the pooling of all the different PCR products before library preparation. To summarize,
195 we performed PCRs on 51 samples, plus 3 extraction controls, plus 5 negative and 2
196 positive PCR controls (total: 61 samples including controls, arranged in one 96-wells
197 micrititer plate; 61 samples X 8 PCRs = 488 PCR reactions). Sequencing was performed
198 by 2 x 125 bp pair-end sequencing on an Illumina HiSeq 2500 platform, which returned
199 a total of 8'321'810 reads.
200
201 **Sequences filtering and taxa assignment**
202 DNA sequences were filtered using OBITools software (Boyer et al., 2016), following the
203 protocol described in Pansu et al. (2015). The obtained sequences were then assigned to
204 the relevant taxa using the *ecotag* program that locates highly similar sequence(s) in a
205 suitable database. Here, sequences were compared to a global plant database generated
206 using silico PCR from EMBL with the *ecoPCR* program (Ficetola et al., 2010) and a BLAST
207 search was conducted using NCBI (http://blast.ncbi.nlm.nih.gov) for taxon attribution.

208 Only sequences with a similarity of >90% to a known taxon were kept for subsequent
209 analyses.
210 In order to remove potential contaminants and PCR errors, we adopted the following
211 conservative selection procedure. Firstly, all sequences identified to a taxa (species,
212 group of species or family) currently absent from Benin were discarded. Secondly, we
213 considered a sequence genuine in a PCR product if its count was > 10 reads. Thirdly, we
214 only kept sequences that reached this threshold in at least one replica in the core
215 samples, and that did not reach this threshold in any of the control replicas; we then
216 grouped sequences with the same identification (Parducci et al., 2017).
217

218 **Geochemical and sedimentological analyses**

219 All geochemical and sedimentological analyses were performed on the half intact cores
220 collected with the Livingston corer. In order to characterize the sediment from which
221 *sed*DNA was extracted, Loss On Ignition (LOI) analyses were performed on 1 cm^3 of
222 sediment, following the procedure described by Heiri et al. (2001). In total, 48 samples
223 were collected from the Livingston core at the depth levels of the *sed*DNA analysis
224 excluding the very upper layers, where only one sample was collected at 8.5 cm, which
225 almost correspond to samples collected at a 8.8, 10.4, 11.7 cm depth for *sed*DNA.
226 Continuous X-ray fluorescence (XRF) analyses were performed with a core scanner
227 Avaatech (X-Ray beam generated with a rhodium anode) at EDYTEM Laboratory (CNRS-
228 University Savoie Mont Blanc, France). Two runs of analyses were performed with a 5
229 mm resolution. For the first run a voltage of 10 kV, an intensity of 1.2 mA and a counting
230 time of 15 s were applied; for the second run, a voltage of 30 kV, an intensity of 0.75 mA
231 and a counting time of 30 s were applied.
232

233 **Pollen assemblages**

234 The pollen counts used for comparison with *sed*DNA were previously published by
235 Salzmann and Hoelzmann (2005). Our sediment core was collected from the same part
236 of the lake. The age-depth model was recalculated from Salzmann and Hoelzmann
237 (2005) using Bacon software (Blaauw and Christen, 2011).

238

239 **Results**

240 ***Plant* DNA *extraction and identification***

241 *Sed*DNA was successfully amplified and sequenced from 33 of the initial 50 samples.
242 Sixteen samples were lost during extraction using the NucleoSpin® Soil kit because fine
243 sediments clogged the kit filters, hampering the washing and elution steps of the
244 extraction*.* One sample was sequenced but without DNA being recovered.
245 A total of 5,560,355 raw reads could be assigned to the 33 samples (Appendix S1). After
246 cleaning with *obiclean*, 4,971,763 reads and 5,365 unique sequences remained. When
247 we retained only those sequences with a similarity of >90% to a known taxon, we
248 obtained 1,443,822 reads and 271 unique sequences (Appendix S2). From among these
249 sequences, we aggregated those attributable to the same high-level taxon (e.g. family);
250 these sequences were checked again by a BLAST search. After aggregation, 148 taxa
251 were identified, of which 74 were considered as exotic (compatible with no species
252 present today in Benin) and 30 were discarded because they were either detected in
253 controls or not significantly present (< 10 reads) in only one replica. This left 43 single
254 taxa. The exotic sequences not in the control but removed were assigned to Betulaceae,
255 *Alnus*, *Cedrus*, *Pelargonium*, *Cardamine*, *Holcus*, *Prunus,* and *Pinaceae*. None of these taxa
256 are native to tropical lowlands and their presence in the samples might be due to
257 contamination. Taxa present in controls were Solanaceae, Lamiaceae, Poaceae, *Musa*,

258 Apiaceae, *Crassula*, Polygonaceae, Amaranthaceae, Rubiaceae, and Fabaceae, and were
259 therefore removed from analyses.
260 The number of taxa and reads replicates decreases with the sediment's age (Figure 2).
261 The first six samples, corresponding to the first 25 cm depth (probably less than 300
262 years old), presented the clearest evidence of PCR success with several taxa detected,
263 each in seven or eight PCR replicates. In the deepest sections of the core the success of
264 DNA amplification was much lower, with the majority of samples being given a taxon
265 detected in only one PCR (figure 3).
266
267 Among the 43 taxa, 22 were identified at the family level and 21 at the genus level. The
268 species level was not considered because of the absence of an accurate local reference
269 library for taxonomic assignation. However, the herb *Sphenoclea zeylanica*, which occurs
270 in damp habitats throughout the tropics, is an exception because the sequences of two
271 species of the genus were available in the GenBank.
272 The LOI550 represents the quantity of organic matter in the sediment (% of the dry
273 weight) that was of aquatic and/or terrestrial origin (Figure 2). Organic matter
274 production induces modifications in the oxygenation conditions at the lake bottom,
275 which also affects the conditions of *sed*DNA preservation. In fact, low oxygen content
276 limits the DNA damage by oxidation and microbial activity, supposedly the primary
277 factor of extracellular DNA degradation through the production of DNase (Blum et al.,
278 1997; Corinaldesi et al., 2005; Willerslev and Cooper, 2005; Parducci et al., 2017). Along
279 the core, the organic matter concentration varied from 11% to 17%. The zirconium
280 coming from zircon (Zr) was generally more concentrated in the sand sediment fraction,
281 and the rubidium (Rb) more enriched in clays (Davies et al., 2015). The ratio Zr/Rb is
282 thus used to trace the changes in the contribution of clay-size particles relative to sand,

283 which can be interpreted as qualitative changes in the erosion dynamic. Titanium (Ti)
284 can also be interpreted as a proxy of runoff and, therefore, rainfall (Metcalfe et al., 2010).
285 The results show short phases with low-erosion rates after the hiatus, during a long
286 period of high-erosive sediment influx from the hiatus of sedimentation until 2,100
287 years cal. BP. Rainfall appears to have decreased until 900 years cal. BP, evident in a
288 more or less constant fine sedimentation due to this lower rainfall. The pollen
289 assemblages presented in Salzmann and Hoelzmann (2005) also confirm a high lake
290 level between 3,000 years cal. BP and 1,000 years cal. BP.
291
292 The comparison with the pollen spectra showed important differences among the
293 diversity and the presence of the main taxa (Figure 3). Some taxa, such as *Nymphaea* and
294 Asteraceae, presented coherent signals. More taxa were recovered from pollen data than
295 from *sed*DNA. However, 13 DNA taxa have no corresponding pollen and spore taxa
296 (Figure 4) such as *Sphenoclea zeylanica*, *Pistia*, or *Ceratophyllum*. *Azolla* is a fern without
297 identifiable spores and *Cosmarium* a green alga without pollen or spores. The dynamic
298 of the *sed*DNA taxa cannot be directly compared with pollen assemblages dynamic
299 because of the scarcity of DNA taxa in the deep part of the core.
300
301
302 **Discussion**
303 1- Lessons for *sed*DNA analyses in lowland tropical lakes
304 For the first time, we obtained amplification of up to 5,000-year-old plant *sed*DNA from
305 lowland tropical lake sediments. While the use of plant *sed*DNA as a palaeoecological
306 proxy has created increasing interest over the last few years – with a focus on mid- to
307 high-latitude lake sediments – its application in the humid tropics has remained limited

308 and focused on high altitudes (Boessenkool et al., 2014) where the temperatures are
309 low, or on very recent sediments (Stoof-Leichsenring et al., 2012). The successful
310 amplification of plant *sed*DNA in lowland tropical lake sediments presented here,
311 therefore, demonstrates the usefulness of the method in low latitude and altitude sites.
312 Our results also emphasize a series of insights on the requirements to retrieve *sed*DNA
313 in general, as well as in the specific context of lowland tropical lakes, namely: i)
314 deposition plant DNA in lake sediments, through taphonomical processes; ii) DNA
315 preservation through time in lake sediments; and iii) technical methods to extract DNA
316 from sediments.
317 Plant *sed*DNA can originate from lacustrine/shoreline plant remains, which readily
318 deposit in the lake, or from terrestrial plant tissues. In some lakes, plants nearby the
319 lakeshore are the main source of a terrestrial plant's DNA (Anderson-Carpenter et al.,
320 2011; Inger Alsos, pers. com). Extracellular DNA from terrestrial plants is, supposedly,
321 bound mainly to soil particles and then transferred and deposited in the lake during soil
322 erosion events (Giguet-Covex et al.; Parducci et al., 2017). In Lake Sele, lacustrine and
323 shoreline sources also dominate the records, even in the most recent sediments. Very
324 few DNA taxa were from terrestrial plants, probably because the majority of DNA from
325 living plant tissue is lost before it reaches the lake sediments. Rapid DNA degradation in
326 a warm tropical climate is a tentative explanation for the limited amount of terrestrial
327 plant DNA successfully amplified. Studies on the soil surfaces of tropical regions have
328 proved that DNA can be preserved in these type of soils (Yoccoz et al., 2012), and in high
329 enough rates that they could be transported to the lake during rainfalls events; however,
330 measures of the degradation rate of this environmental DNA are lacking. An alternative
331 explanation for the absence of DNA preservation is related to the lake-catchment
332 features. Its flat topography and the riverine vegetation likely limit the flow of terrestrial

333 plant DNA from the surrounding soils to the sediments in the middle of the lake. The
334 question of whether DNA is preserved over long periods in sediments from warm
335 climates is subject to controversy. As commonly observed (Anderson-Carpenter et al.,
336 2011; Boessenkool et al., 2014), the number of taxa detected through *sed*DNA and the
337 number of reads tended to decline along the core. Nevertheless, some taxa were
338 detected in sediments even older than 4,500 years. Despite the high tropical
339 temperatures, *sed*DNA was detected. This success reflects the presence of favorable
340 conditions for DNA preservation, which might be due to low ventilation minimizing the
341 oxidation as well as the bacterial activity. Lake Sele is a shallow lake with a depth of less
342 than 1.5m during the dry season, but it is less susceptible to desiccation because of its
343 link to the large groundwater table of the Ouémé River (Adam and Boko, 1983). This
344 would explain why the lake has never experienced any dry out during the last 3,000
345 years. In the Kenyan Lake Naivasha (Stoof-Leichsenring et al., 2012), with a comparable
346 sediment and mean annual temperature, the organic matter content was higher: 20 to
347 50 % in recent sediments (Mergeay et al., 2004). A high quantity of organic matter can
348 induce anoxic conditions, with a low pH, creating unfavorable conditions for the
349 preservation of *sed*DNA, even though it can contain more plant DNA susceptible to
350 preservation over longer time periods. At Lake Sele, it appears that more than 14%
351 organic matter in the sediment is required to allow extraction of *sed*DNA.
352 The specific alluvial nature of sediments (after the hiatus to 900 years. cal. BP) with
353 coarse detrital input during a high lake lev@el (Salzmann and Hoelzmann, 2005), might
354 explain the amount of failures during the purification and extraction processes. Out of
355 51 samples, 16 clogged the kit filters and most replicates of 3 additional samples were
356 found with no, or negligible, DNA. This emphasizes that the extraction and methodology
357 used to purify and extract the *sed*DNA (Pansu et al., 2015) should be adapted to this kind

358 of alluvium sediment. In the samples where extraction was successful, the majority of
359 the eight amplification replicates contained DNA, reinforcing the idea that DNA
360 preservation was good all along the sediment core.
361
362 2- Taxonomic resolution of taxa identified by DNA
363 *Taphonomy of sedDNA*
364 As previously mentioned, the taxa identified by DNA analyses are mainly aquatics plants.
365 This is the case for *Nymphaea* which is well recorded in the upper part of the core with
366 the highest number of reads and PCR numbers; *Azolla, Ludwigia, Persicaria,* and
367 *Sphenoclea zeylanica* that were recorded with the highest reads and PCR numbers are
368 also aquatic plants. The limited depth of the lake also probably favors the abundance of
369 aquatic plants and prevents the development of trees or shrubs nearby, due to the
370 flooding of the shorelines during the wet season. These water oscillations may explain
371 why the amount of *sed*DNA from trees is extremely limited. More taxa were recovered
372 from pollen data than from *sed*DNA; however, some DNA taxa have no corresponding
373 pollen taxa. The aquatic plant *Azolla* is a fern, which probably do not produce
374 recognizable spores, but *Sphenoclea*, with identifiable pollen, should have been
375 recorded. Pollen and *sed*DNA records often show low overlapping (Jørgensen et al.,
376 2012; Parducci et al., 2013; Parducci et al., 2017), which can be explained by the
377 differences in source productions, the taxonomic resolution, and the primers used for
378 sequencing (Parducci et al., 2013). Furthermore, the pollen of some species can be
379 dispersed over long distances, thus pollens can also represent regional vegetation, while
380 *sed*DNA provides a more local signal (Boessenkool et al., 2014; Parducci et al., 2017). For
381 the moment, *sed*DNA data are more related to the presence/absence of taxa, as
382 macroremain analysis are interpreted in paleoecological studies (Birks, 2001). But some

383 recent studies on soils suggest that the number of reads can be related to biomass plant
384 production (Yoccoz et al., 2012; Pansu et al., 2015). In our data, it appears that for some
385 taxa (e.g. *Nymphaea*) the number of reads and pollen percentages actually correlate.
386 Combining pollen and *sed*DNA should help improve vegetation reconstruction, notably
387 in tropical forests where a large amount of trees do not have an entomophilous pollen
388 dispersal mode.
389
390 *Taxonomic reference library*
391 The lack of a complete DNA reference library of local species is a major issue for the
392 application of *sed*DNA for vegetation reconstruction. For instance, one sequence
393 detected matched well (97% identity) with the GenBank sequences of the genus
394 *Pontederia*. As this genus is not currently present in Benin, it is likely that the DNA
395 sequence belongs to *Eichhornia crassipes* (Pontederiaceae), for which no sequences are
396 available in GenBank. This plant was introduced from South America several decades
397 ago and now completely covers the periphery of the lake, which confirms that the DNA
398 signal is essentially coming from the vegetation that is developing on the shoreline and
399 in the shallow parts of the lake.
400 It is also possible that some locally present species, for which no sequences are available
401 in GenBank, were excluded because they matched related taxa with similar sequences
402 that are absent in Benin. Indeed, a large number of exotic taxa exotic from the study area
403 that were found in the lake samples were not from the control samples. The
404 identification of taxa through metabarcoding is severely limited by the availability of
405 sequences within reference databases; and if no sequences of a given taxon are
406 available, taxonomic assignment tools tend to assign sequences to the closest taxon that
407 can be available further apart. The development of extensive reference databases is,

408 therefore, a key priority to fully exploit the power of metabarcoding for environmental
409 reconstruction (Taberlet et al., 2012a). Efforts to map the genetic diversity of vegetation
410 are ongoing; unfortunately these mainly focus on Europe and North America, while
411 information on highly diverse tropical areas remains scarce (Coissac et al., 2016;
412 Miraldo et al., 2016).
413
414 3. Agriculture recorded by *sed*DNA
415 We detected the DNA of the genus *Ipomoea* (Convolvulaceae) in recent sediments.
416 Although the specific attribution is not certain, it is likely that the DNA signal was
417 produced by sweet potatoes (*Ipomoea batatas*); these were introduced to the Lake Sele
418 region from South America, probably at the same time as *Eichhornia crassipes,* in the
419 sixteenth century (O'Brien, 1972). There is little chance that another *Ipomoea* may have
420 produced this DNA signal, given that the pollen record indicates a continuous presence
421 of (non-cultivated) Convolvulaceae growing presumably on the lakeshores since the
422 level dropped after 3,000 cal. BP.
423 The original aim of this study was to test if agriculture could be recorded by *sed*DNA.
424 The detection of *Ipomoeae sp.* (sweet potato) DNA in recent sediments suggests that
425 *sed*DNA has the potential to trace the history of agricultural developments in these
426 environments. Nowadays sweet potato is cultivated closed to permanent water, on the
427 shoreline during the dry season when the lake is low. In some lakes, the DNA of plants
428 growing near the shore is more easily detected than the DNA of plants living further
429 away (Inger Alsos, pers. com.), which may explain why this species is better recorded
430 than the cassava which is grown away from the shore. However, it is surprising that
431 cassava or even Euphorbiaceae were not detected in the recent sediment. One possible

432 explanation is that parts of the cultivated plant, after gathering tubers and fresh leaves,
433 are taken and burnt somewhere that prevents possible transfer to the lake sediments.
434 We also found sporadic contamination by food crops such as the banana or cacao. DNA
435 from *Musa* was detected in several samples along the core but also in the controls. Even
436 more problematic are amplified sequences matching with European trees (*Cedrus*,
437 *Betula* and Pinaceae), which were detected in the samples but not in the control. This
438 highlights the complexity of using ancient DNA for the study of domestic and cultivated
439 species (Weiß et al., 2015). The controversial study of Smith et al. (2015) concerning
440 wheat cultivated 8,000 years ago in the British Isles, before the Neolithic introduction of
441 domestic cereals, is an illustrative example.
442

443 **Conclusion**

444 Our results confirm the usefulness of *sed*DNA for the study of tropical sediments, as DNA
445 can be preserved for thousand of years even in "hot" environmental conditions. The
446 preservation seems to decrease with age but the quality of the sediment may be a more
447 important factor for the successful analysis of *sed*DNA. Fine clayous sediments are more
448 susceptible to contain extractable *sed*DNA and a minimum of organic matter in the bulk
449 sediment is necessary. A small lake catchment with slopes limiting the development of
450 large riverine aquatic plants is also probably a better configuration to record a DNA
451 terrestrial signature than a flat lake catchment. *Sed*DNA is a valuable approach in
452 tropical areas for searching specific plants growing close to the water. The plant
453 diversity recovered through *sed*DNA in Lake Sele clearly provides a local signal. Volcano
454 crater lakes with inner ring vegetation, like most of the East African lakes, would be
455 ideal to attempt new *sed*DNA studies. Also, although mammal DNA was not researched

456 in our sediments it must be considered because in tropical areas, more than anywhere
457 else, permanent water is an attractive place for wild fauna and domestic stocks.
458
459 **Acknowledgment**
460 This work was supported by project PEPS-CNRS "Banana Connexion". We would like to
461 thank Benoit Brossier for his technical support during field trip. We also thank Allowen
462 Evin for fruitful discussions and Frederic Boyer for raw data handling. We thank Jill
463 Cucchi for professional English editing.
464
465 **Conflict of interest**
466 The authors would like to mention that LG is one of the co-inventors of patents related
467 to g-h primers and the subsequent use of the P6 loop of the chloroplast *trn*L (UAA)
468 intron for plant identification using degraded template DNA. These patents only restrict
469 commercial applications and have no impact on the use of this locus by academic
470 researchers.
471
472
473
474
475
476
477
478
479
480

481 **References**

482 1 Adam, K.S., Boko, M., 1983. Le Bénin. les Ed. du Flamboyant EDICEF, Paris.
483
484 2 Allentoft, M.E., Collins, M., Harker, D., Haile, J., Oskam, C.L., Hale, M.L., Campos,
485 P.F., Samaniego, J.A., Gilbert, M.T.P., Willerslev, E., 2012. The half-life of DNA in
486 bone: measuring decay kinetics in 158 dated fossils, Proc. R. Soc. B. The Royal
487 Society, pp. 4724-4733.
488
489 3 Alsos, I.G., Sjögren, P., Edwards, M.E., Landvik, J.Y., Gielly, L., Forwick, M., Coissac,
490 E., Brown, A.G., Jakobsen, L.V., Føreid, M.K., 2016. Sedimentary ancient DNA from
491 Lake Skartjørna, Svalbard: Assessing the resilience of arctic flora to Holocene
492 climate change. The Holocene 26, 627-642.
493
494 4 Anderson-Carpenter, L.L., McLachlan, J.S., Jackson, S.T., Kuch, M., Lumibao, C.Y.,
495 Poinar, H.N., 2011. Ancient DNA from lake sediments: Bridging the gap between
496 paleoecology and genetics. BMC Evolutionary biology 11:30, 1-15.
497
498 5 Binladen, J., Gilbert, M.T.P., Bollback, J.P., Panitz, F., Bendixen, C., Nielsen, R.,
499 Willerslev, E., 2007. The use of coded PCR primers enables high-throughput
500 sequencing of multiple homolog amplification products by 454 parallel
501 sequencing. PLoS ONE 2, e197.
502
503 6 Birks, H.H., 2001. Plant macrofossils, In: Smol, J., Birks, H., Last, W. (Eds.),
504 Terrestrial algal and siliceous indicators. Tracking environmental change using
505 lake sediments. Kluwer Academic Publishers, Dordrecht, pp. 49-74.
506
507 7 Blaauw, M., 2010. Methods and code for 'classical'age-modelling of radiocarbon
508 sequences. Quaternary Geochronology 5, 512-518.
509
510 8 Blaauw, M., Christen, J.A., 2011. Flexible paleoclimate age-depth models using an
511 autoregressive gamma process. Bayesian Analysis 6, 457-474.
512
513 9 Blum, S.A.E., Lorenz, M.G., Wackernagel, W., 1997. Mechanism of Retarded DNA
514 Degradation and Prokaryotic Origin of DNases in Nonsterile Soils. Systematic and
515 Applied Microbiology 20, 513-521.
516
517 10 Boessenkool, S., McGlynn, G., Epp, L.S., Taylor, D., Pimentel, M., Gizaw, A.,
518 Nemomissa, S., Brochmann, C., Popp, M., 2014. Use of Ancient Sedimentary DNA
519 as a Novel Conservation Tool for High-Altitude Tropical Biodiversity.
520 Conservation Biology 28, 446-455.
521
522 11 Born, C., Kjellberg, F., Chevallier, M.-H., Vignes, H., Dikangadissi, J.-T., Sanguié, J.,
523 Wickings, E.J., Hossaert-McKey, M., 2008. Colonization processes and the
524 maintenance of genetic diversity: insights from a pioneer rainforest tree,
525 Aucoumea klaineana. Proceedings of the Royal Society B: Biological Sciences 275,
526 2171-2179.
527

528 12 Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., Coissac, E., 2016. obitools:
529 a unix - inspired software package for DNA metabarcoding. Molecular Ecology
530 Resources 16, 176-182.
531
532 13 Britton, J.R., Boar, R., Grey, J., Foster, J., Lugonzo, J., Harper, D., 2007. From
533 introduction to fishery dominance: the initial impacts of the invasive carp
534 Cyprinus carpio in Lake Naivasha, Kenya, 1999 to 2006. Journal of Fish Biology
535 71, 239-257.
536
537 14 Coissac, E., Hollingsworth, P.M., Lavergne, S., Taberlet, P., 2016. From barcodes to
538 genomes: extending the concept of DNA barcoding. Molecular Ecology.
539
540 15 Corinaldesi, C., Danovaro, R., Dell'Anno, A., 2005. Simultaneous recovery of
541 extracellular and intracellular DNA suitable for molecular studies from marine
542 sediments. Appl Environ Microbiol 71, 46-50.
543
544 16 Cuni-Sanchez, A., White, L.J., Calders, K., Jeffery, K.J., Abernethy, K., Burt, A.,
545 Disney, M., Gilpin, M., Gomez-Dans, J.L., Lewis, S.L., 2016. African Savanna-Forest
546 Boundary Dynamics: A 20-Year Study. PLoS ONE 11, e0156934.
547
548 17 Davies, S.J., Lamb, H.F., Roberts, S.J., 2015. Micro-XRF core scanning in
549 palaeolimnology: Recent developments, Micro-XRF Studies of Sediment Cores.
550 Springer, pp. 189-226.
551
552 18 Epp, L.S., Stoof, K.R., Trauth, M.H., Tiedemann, R., 2010. Historical genetics on a
553 sediment core from a Kenyan lake: intraspecific genotype turnover in a tropical
554 rotifer is related to past environmental changes. Journal of Paleolimnology 43,
555 939-954.
556
557 19 Ficetola, G.F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P.,
558 Pompanon, F., 2010. An in silico approach for the evaluation of DNA barcodes.
559 BMC genomics 11, 1.
560
561 20 Ficetola, G.F., Pansu, J., Bonin, A., Coissac, E., Giguet - Covex, C., De Barba, M.,
562 Gielly, L., Lopes, C.M., Boyer, F., Pompanon, F., 2015. Replication levels, false
563 presences and the estimation of the presence/absence from eDNA
564 metabarcoding data. Molecular Ecology Resources 15, 543-556.
565
566 21 Fuller, D.Q., Denham, T., Arroyo-Kalin, M., Lucas, L., Stevens, C.J., Qin, L., Allaby,
567 R.G., Purugganan, M.D., 2014. Convergent evolution and parallelism in plant
568 domestication revealed by an expanding archaeological record. Proceedings of
569 the National Academy of Sciences 111, 6147-6152.
570
571 22 Giguet-Covex, C., Ficetola, G.F., Walsh, K.J., Poulenard, J., Bajard, M., Fouinat, L.,
572 Sabatier, P., Gielly, L., David, F., Taberlet, P., Arnaud, F., New insights on lake
573 sediment DNA from the catchment: importance of taphonomic and analytical
574 issues on the record quality. submitted QSR.
575

576 23 Giguet-Covex, C., Pansu, J., Arnaud, F., Rey, P.-J.r.m., Griggo, C., Gielly, L., Domaizon,
577 I., Coissac, E., David, F., Choler, P., Poulenard, J.r.m., Taberlet, P., 2014. Long
578 livestock farming history and human landscape shaping revealed by lake
579 sediment DNA. Nat Commun 5.
580
581 24 Haile, J., Froese, D.G., MacPhee, R.D.E., Roberts, R.G., Arnold, L.J., Reyes, A.V.,
582 Rasmussen, M., Nielsen, R., Brook, B.W., Robinson, S., Demuro, M., Gilbert, M.T.P.,
583 Munch, K., Austin, J.J., Cooper, A., Barnes, I., Möller, P., Willerslev, E., 2009. Ancient
584 DNA reveals late survival of mammoth and horse in interior Alaska. Proceedings
585 of the National Academy of Sciences 106, 22352-22357.
586
587 25 Heiri, O., Lotter, A.F., Lemcke, G., 2001. Loss on ignition as a method for
588 estimating organic and carbonate content in sediments: reproducibility and
589 comparability of results. Journal of Paleolimnology 25, 101-110.
590
591 26 Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2004. The WorldClim
592 interpolated global terrestrial climate surfaces. Version 1.3.
593
594 27 Jørgensen, T., Haile, J., Möller, P., Andreev, A., Boessenkool, S., Rasmussen, M.,
595 Kienast, F., Coissac, E., Taberlet, P., Brochmann, C., 2012. A comparative study of
596 ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of
597 northern Siberia reveals long - term vegetational stability. Molecular Ecology 21,
598 1989-2003.
599
600 28 Matisoo-Smith, E., Roberts, K., Welikala, N., Tannock, G., Chester, P., Feek, D.,
601 Flenley, J., 2008. Recovery of DNA and pollen from New Zealand lake sediments.
602 Quaternary International 184, 139-149.
603
604 29 Mergeay, J., Verschuren, D., Kerckhoven, L.V., Meester, L.D., 2004. Two hundred
605 years of a diverse Daphnia community in Lake Naivasha (Kenya): effects of
606 natural and human-induced environmental changes. Freshwater Biology 49, 998-
607 1013.
608
609 30 Metcalfe, S.E., Jones, M.D., Davies, S.J., Noren, A., MacKenzie, A., 2010. Climate
610 variability over the last two millennia in the North American Monsoon region,
611 recorded in laminated lake sediments from Laguna de Juanacatlán, Mexico. The
612 Holocene 20, 1195-1206.
613
614 31 Miraldo, A., Li, S., Borregaard, M.K., Flórez-Rodríguez, A., Gopalakrishnan, S.,
615 Rizvanovic, M., Wang, Z., Rahbek, C., Marske, K.A., Nogués-Bravo, D., 2016. An
616 Anthropocene map of genetic diversity. Science 353, 1532-1535.
617
618 32 O'Brien, P.J., 1972. The sweet potato: its origin and dispersal. American
619 Anthropologist 74, 342-365.
620
621 33 Pansu, J., Giguet - Covex, C., Ficetola, G.F., Gielly, L., Boyer, F., Zinger, L., Arnaud, F.,
622 Poulenard, J., Taberlet, P., Choler, P., 2015. Reconstructing long - term human
623 impacts on plant communities: an ecological approach based on lake sediment
624 DNA. Molecular Ecology 24, 1485-1498.

625
626 34 Parducci, L., Bennet, K.D., Ficetola, G.F., Alsos, I.G., Suyama, Y., Wood, J.R.,
627 Pedersen, M.W., 2017. Ancient plant DNA in lake sediments. New Phytologist .
628 DOI: 10.1111/nph.14470.
629
630 35 Parducci, L., Jørgensen, T., Tollefsrud, M.M., Elverland, E., Alm, T., Fontana, S.L.,
631 Bennett, K.D., Haile, J., Matetovici, I., Suyama, Y., Edwards, M.E., Andersen, K.,
632 Rasmussen, M., Boessenkool, S., Coissac, E., Brochmann, C., Taberlet, P., Houmark-
633 Nielsen, M., Larsen, N.K., Orlando, L., Gilbert, M.T.P., Kjær, K.H., Alsos, I.G.,
634 Willerslev, E., 2012. Glacial Survival of Boreal Trees in Northern Scandinavia.
635 Science 335, 1083-1086.
636
637 36 Parducci, L., Matetovici, I., Fontana, S.L., Bennett, K.D., Suyama, Y., Haile, J., Kjær,
638 K.H., Larsen, N.K., Drouzas, A.D., Willerslev, E., 2013. Molecular- and pollen-based
639 vegetation analysis in lake sediments from central Scandinavia. Molecular
640 Ecology 22, 3511-3524.
641
642 37 Pedersen, M.W., Ginolhac, A., Orlando, L., Olsen, J., Andersen, K., Holm, J., Funder,
643 S., Willerslev, E., Kjaer, K.H., 2013. A comparative study of ancient environmental
644 DNA to pollen and macrofossils from lake sediments reveals taxonomic overlap
645 and additional plant taxa. Quaternary Science Reviews 75, 161-168.
646
647 38 Salzmann, U., Hoelzmann, P., 2005. The Dahomey Gap: an abrupt climatically
648 induced rain forest fragmentation in West Africa during the late Holocene. The
649 Holocene 15, 190-199.
650
651 39 Smith, O., Momber, G., Bates, R., Garwood, P., Fitch, S., Pallen, M., Gaffney, V.,
652 Allaby, R.G., 2015. Sedimentary DNA from a submerged site reveals wheat in the
653 British Isles 8000 years ago. Science 347, 998-1001.
654
655 40 Stoof-Leichsenring, K.R., Epp, L.S., Trauth, M.H., Tiedemann, R., 2012. Hidden
656 diversity in diatoms of Kenyan Lake Naivasha: a genetic approach detects
657 temporal variation. Molecular Ecology 21, 1918-1930.
658
659 41 Stoof-Leichsenring, K.R., Junginger, A., Olaka, L.A., Tiedemann, R., Trauth, M.H.,
660 2011. Environmental variability in Lake Naivasha, Kenya, over the last two
661 centuries. Journal of Paleolimnology 45, 353-367.
662
663 42 Taberlet, P., Coissac, E., Hajibabaei, M., Rieseberg, L.H., 2012a. Environmental
664 DNA. Molecular Ecology 21, 1789-1793.
665
666 43 Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat,
667 T., Corthier, G., Brochmann, C., Willerslev, E., 2007. Power and limitations of the
668 chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic acids research 35,
669 e14.
670
671 44 Taberlet, P., Prud'Homme, S.M., Campione, E., Roy, J., Miquel, C., Shehzad, W.,
672 Gielly, L., Rioux, D., Choler, P., Clément, J.-C., 2012b. Soil sampling and isolation of

673 extracellular DNA from large amount of starting material suitable for
674 metabarcoding studies. Molecular Ecology 21, 1816-1820.
675
676 45 Thomsen, P.F., Willerslev, E., 2015. Environmental DNA – An emerging tool in
677 conservation for monitoring past and present biodiversity. Biological
678 Conservation 183, 4-18.
679
680 46 Valentini, A., Miquel, C., Nawaz, M.A., Bellemain, E., Coissac, E., Pompanon, F.,
681 Gielly, L., Cruaud, C., Nascetti, G., Wincker, P., 2009. New perspectives in diet
682 analysis based on DNA barcoding and parallel pyrosequencing: the trnL
683 approach. Molecular Ecology Resources 9, 51-60.
684
685 47 Weiß, C.L., Dannemann, M., Prüfer, K., Burbano, H.A., 2015. Contesting the
686 presence of wheat in the British Isles 8,000 years ago by assessing ancient DNA
687 authenticity from low-coverage data. eLife 4, e10005.
688
689 48 White, F., 1983. The vegetation of Africa, a descriptive memoir to accompany the
690 UNESCO/AETFAT/UNSO vegetation map of Africa. Unesco, Paris.
691
692 49 Willerslev, E., Cooper, A., 2005. Ancient DNA. Proc Biol Sci 272, 3-16.
693
694 50 Willerslev, E., Hansen, A.J., Binladen, J., Brand, T.B., Gilbert, M.T.P., Shapiro, B.,
695 Bunce, M., Wiuf, C., Gilichinsky, D.A., Cooper, A., 2003. Diverse Plant and Animal
696 Genetic Records from Holocene and Pleistocene Sediments. Science 300, 791-
697 795.
698
699 51 Yoccoz, N., Bråthen, K., Gielly, L., Haile, J., Edwards, M., Goslar, T., Von Stedingk, H.,
700 Brysting, A., Coissac, E., Pompanon, F., 2012. DNA from soil mirrors plant
701 taxonomic and growth form diversity. Molecular Ecology 21, 3647-3655.
702
703
704
705
706

Figure captions

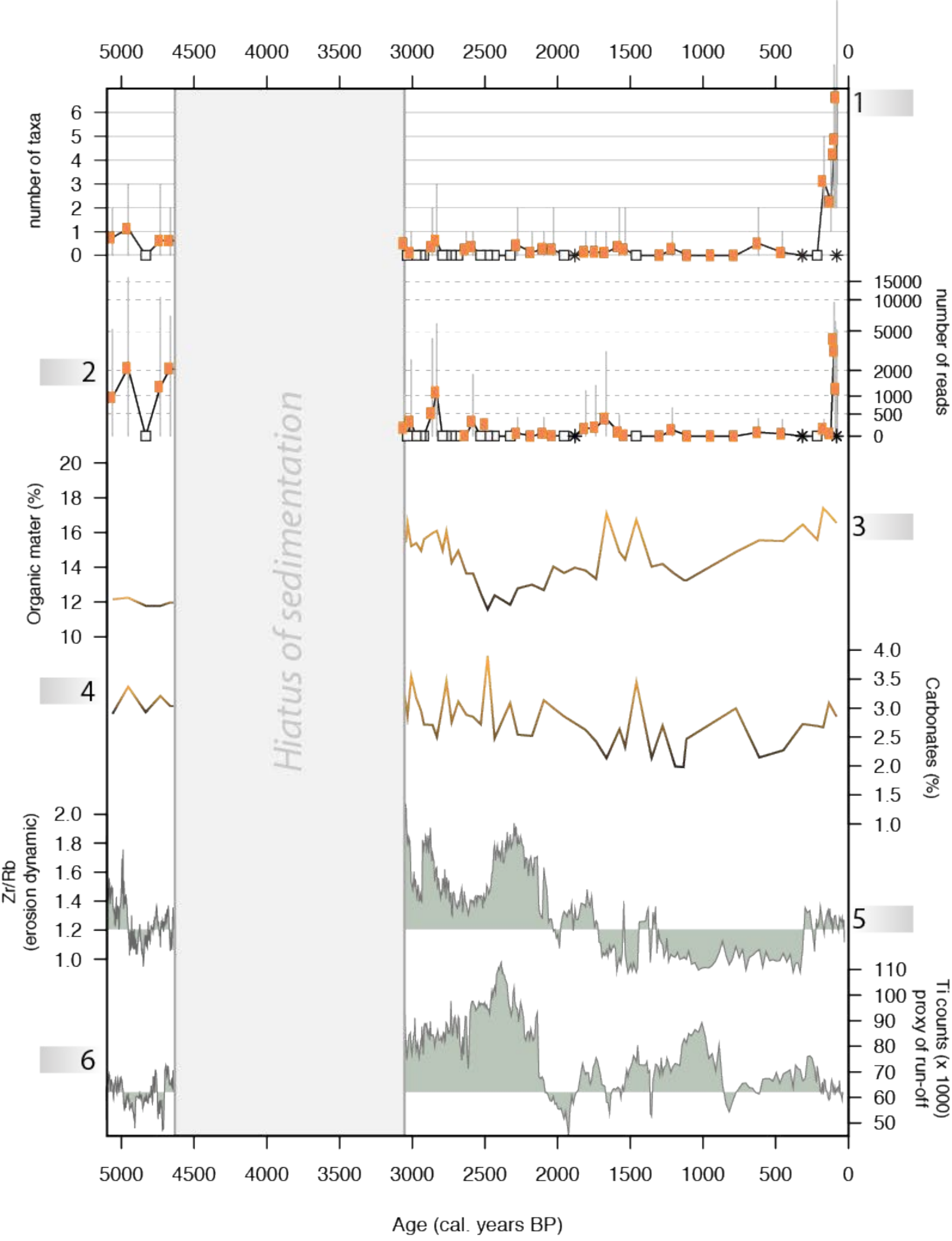
Figure 1. Localization of the coring site in Lake Sele (Benin). The annually submerged shorelines are identifiable by the greenness of the aquatic plants and the agricultural parcels which present geometric forms due to the shapes of the fields.

Figure 2. Comparison between *sed*DNA and sedimentology during the last 5,000 years in Lake Sele: 1) Number of identified taxa (mean values and standard deviations of the replicates) and 2) number of reads (mean values and standard deviations of the replicates). The stars represent the samples where PCR did not permit extraction of DNA, and the white squares are the samples that clogged the kit filters and hampered extraction. 3) Organic matter content ($LOI_{550°C}$). 4) Carbonate content ($LOI_{950°C}$). 5) Zr/R–b as proxy of erosion dynamic (high values = coarser grain-size), and 6) Ti related to increased run-off/rainfall.

Figure 3. Comparison between the *sed*DNA of plants (in orange) and pollen data (in green) from Lake Sele. The DNA signal is expressed as a percentage of the total number of replicates per level. The size of the circles indicates the number of reads for the taxon. The pollen percentage is calculated using the total pollen sum excluding undetermined, modified from Salzmann and Hoelzmann (2005).

Figure 4. Number of taxa identified, per 200 years intervals, by *sed*DNA (orange) and pollen (green). Common taxa are in dark orange.

N
SELE-1&2
500m
Image © 2016 CNES / Astrium
© 2016 Google
Google earth



number of taxa
number of reads
Organic mater (%)
Carbonates (%)
Zr/Rb
(erosion dynamic)
Ti counts (x 1000)
proxy of run-off
Hiatus of sedimentation
Age (cal. years BP)
1
2
3
4
5
6

10,000 reads
1,000
100
10

Aquatic plants

Nymphaea
100%
0
Nymphaea
4%
3
2
1
0

Azolla
100%
0
Salviniaceae undiff.
0.4%
0.2
0.0

Ludwigia
100%
0
Ludwigia
0.8%
0.4
0.0

Pontederiaceae cf. Pontederia
100%
0
Eichornea
0.2%
0.1
0.0

Pistia
100%
0

Ceratophyllum
100%
0

Green algae

Cosmarium
100%
0

5000 4000 3000 2000 1000 0
Age (cal. years BP)

riverine plants

Persicaria
100%
0
Persicaria
0.4%
0.2
0.0

Sphenoclea zeylanica
100%
0

Ipomoea (cf. Ipomea batata - Sweet potato)
100%
0
Ipomoea
0.15%
0.0

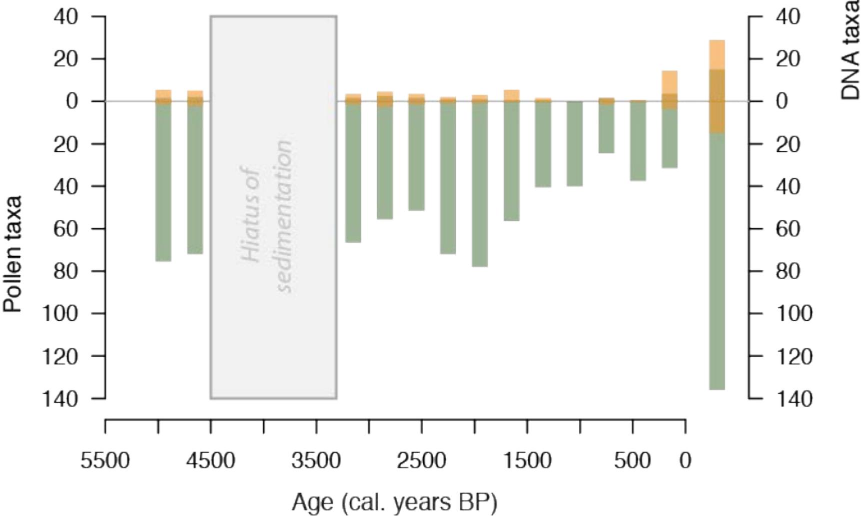
Undif. terrestrial plants

Asteraceae
100%
0
Asteraceae
1.2%
0.8
0.4
0.0

Fabaceae
100%
0
Fabaceae
3.0%
2.0
1.0
0.0

Poaceae PACMAD
100%
0
Poaceae
60%
40
20
0

5000 4000 3000 2000 1000 0
Age (cal. years BP)


Pollen taxa
DNA taxa
Hiatus of sedimentation
Age (cal. years BP)
5500
4500
3500
2500
1500
500
0
40
20
0
20
40
60
80
100
120
140