

PREPRINT

Author-formatted, not peer-reviewed document posted on 27/06/2024

DOI: <https://doi.org/10.3897/arphapreprints.e130574>

DNA metabarcoding, an efficient way to detect non-native cerambycid beetles in trapping collections?

**Lois Veillat,  Stephane Boyer, Marina Querejeta Coma, Emmanuelle Magnoux,  Alain Roques, 
Carlos Lopez-Vaamonde, Géraldine Roux**

1 To be submitted to Neobiota

2

3 **DNA metabarcoding, an efficient way to detect non-native** 4 **cerambycid beetles in trapping collections?**

5

6

7 Loïs Veillat¹, Stéphane Boyer², Marina Querejeta^{2,4}, Emmanuelle Magnoux¹, Alain
8 Roques¹, Carlos Lopez-Vaamonde^{1,2*}, Geraldine Roux^{1,3*}

9

10 ¹INRAE, UR633, Zoologie Forestière, F-45075 Orléans, France

11 ²Institut de Recherche sur la Biologie de l’Insecte, UMR7261 CNRS - Université de Tours,
12 France

13 ³Laboratoire Physiologie, Ecologie et Environnement P2E, Université d’Orléans, 1 rue de
14 Chartres, B.P. 6759, 45067 Orléans cedex 2, France

15 ⁴Department of Functional Biology, University of Oviedo, Oviedo, Asturias, Spain

16

17 Correspondance : Loïs Veillat and Géraldine Roux, INRAE, UR633, Zoologie Forestière, F-
18 45075 Orléans, France. lois.veillat@inrae.fr ; geraldine.roux@univ-orleans.fr

19

20 *These authors contributed equally to this work

21

22

23 **Abstract**

24

25 Individual sorting and identification of thousands of insects collected in mass trapping

26 biosurveillance programs is a labour intensive and time-consuming process.

27 Metabarcoding, allows for the simultaneous identification of multiple individuals in a

28 single mixed sample and has the potential to expedite this process. However, detecting

29 all the species present in a bulk sample can be challenging, especially when under-

30 represented non-native specimens are intercepted.

31

32 In this study, we quantified the effectiveness of DNA metabarcoding at detecting exotic

33 species within six different mock communities of native and non-native species of

34 European xylophagous cerambycid beetles. The main objective is to compare three

35 different sequencing technologies (MinION, Illumina, and IonTorrent) to evaluate which

36 one is the most suitable in this context. Although we did not observe significant
37 differences in the total number of species detected between the three sequencing
38 technologies, MinION detected a greater number of species on field-like samples. All
39 three sequencing technologies achieved in detecting and identifying closely related
40 species and species at low abundance. The capture method of insects in the field greatly
41 influences sample preservation and detection. Individuals captured in traps containing
42 monopropylene and water had lower DNA concentration, leading to lower species
43 detection rates compared to individuals killed using just an insecticide without any
44 collection medium.

45

46 **Keywords:** Alien, Biological invasions, Biosecurity, Cerambycidae, Exotic, Illumina®,
47 IonTorrent®, Oxford Nanopore®, Xylophagous.

48

49 **Introduction**

50 The exponential increase in biological invasions that has been observed over the past
51 decades is expected to persist (Seebens et al. 2021). This is primarily due to factors such
52 as globalization, tourism, and global warming (Chown et al. 2015). Among the species
53 introduced beyond their native range by human activities, insects are the most prevalent
54 group (Seebens et al. 2018) and can cause a wide range of impacts. Non-native insects
55 can affect native flora, fauna and ecosystems in various ways (Kenis et al. 2009) and they
56 can as well transmit pathogens and diseases, thus threatening public health (Mazza et al.
57 2014). Economic implications are also to be considered since numerous invasive insects
58 are important pests for agricultural crops and plantation forests, inducing huge
59 management costs (Bradshaw et al. 2016).

60

61 Among these non-native insects, species associated with woody plants are increasingly
62 dominating, accounting for 76.5% of all herbivore species newly recorded in Europe
63 from 2000 to 2014 potentially because of the growing trade of ornamental plants and
64 wooden packaging material transported in international cargo shipments (Aukema et al.
65 2010; Roques et al. 2016). One of these important families of xylophagous beetles is the
66 long-horned Cerambycidae, with more than 200 species affecting forestry, horticulture,
67 and agriculture (Rossa and Goczał 2021), and resulting in multimillion-dollar losses
68 every year (Wang 2017). To detect potential new invasions of Cerambycids,
69 biomonitoring programs have been set up over large geographical areas with intensive
70 trapping campaigns extending over several years (Roques et al. 2023; Mas et al. 2023).
71 However, rapidly evolving trades lead to changes in trade routes and imported goods
72 which results in an increasing arrival of new non-native species. Many of these species
73 have not been previously reported as invaders, some are not considered to be pests in
74 their native ranges, and some could even be unknown to science (Seebens et al. 2018).
75 As part of the European project HOMED (<https://homed-project.eu/>) 244 Cerambycid
76 traps were set up across Europe (France, Italy, Spain, Switzerland, Portugal, Austria,
77 England, Greece, Slovenia, Netherlands, Bulgaria, Czech Republic, and in Sweden), 38 in
78 Asia (China, Siberia, Russia), 11 in North America (USA, Canada), five in the Caribbean
79 (Martinique) and four in Australia, all baited with generic lures, for simultaneous
80 detection of multiple species (Roques et al. 2023). In such large-scale trapping
81 campaigns, thousands of captured insects must be sorted and identified by expert
82 taxonomists. This identification step is time-consuming and labor-intensive, thus
83 limiting the rapid detection of non-native individuals among large numbers of native

84 ones (Piper et al. 2019; Abeynayake et al. 2021; Chua et al. 2023). Yet, it is essential that
85 those non-native species are detected as quickly as possible to allow their eradication
86 before establishment and dispersal (Richardson et al. 2000; Blackburn et al. 2011;
87 Giovani et al. 2020).

88

89 For insects, traditional DNA barcoding, using a short fragment of the Cytochrome
90 Oxidase 1 (*COI*) gene, has truly become a universal tool to identify a species whatever the
91 life stage (Hebert et al., 2003). Namely, DNA barcoding has been successfully used to
92 accurately identify cerambycid pest species for biomonitoring (Hodgetts et al. 2016, Wu
93 et al. 2017, Kelnarova et al. 2019, Javal et al. 2021). Despite its numerous advantages,
94 individual DNA barcoding remains a laborious and time-consuming approach in the
95 context of mass trapped insects as it requires individual sorting of thousands of
96 specimens, tissue sampling (often legs), DNA extraction and amplification and finally
97 sequencing of each sample individually. However, the recent application of high-
98 throughput sequencing (HTS) technologies to DNA barcoding allows to expedite the
99 production of thousands of DNA barcodes (deWaard et al. 2019; Srivathsan et al. 2021).

100

101 This metabarcoding approach generates a large number of short DNA sequences (reads),
102 allowing the accurate identification of multiple species simultaneously from a single
103 mixed sample (hereafter called “bulk”) (Liu et al. 2020), such as all the individual insects
104 captured in a single biomonitoring trap. Moreover, compared to traditional
105 morphological identification, metabarcoding offers a significant reduction in costs
106 (Batovska et al. 2021), generally providing equivalent or better detection and identifying
107 a much wider spectrum of taxa (Elbrecht et al. 2017; Andújar et al. 2018). Using DNA as

108 a proxy for species detection and considering sequence variation within and among taxa,
109 metabarcoding approaches are however constrained by the completeness of the
110 reference databases to accurately assign sequences to correctly identified taxa (Liu et al.
111 2020).

112

113 Although metabarcoding has several advantages, ensuring the accuracy of detections is
114 crucial. Erroneous detections of pest species can have severe environmental and
115 economic consequences (Batovska et al. 2021). However, metabarcoding approaches still
116 suffer from methodological limitations that may make them unsuitable for rapid
117 biosecurity detection (contaminations, limited quantitative aspect, incomplete
118 databases, false positives, etc.). One specific challenge is the time required to process
119 samples, which can be problematic when there are long delays between capturing
120 individuals and obtaining sequencing results. This is especially true when sampling sites
121 are located far away from laboratories, when transporting samples may require specific
122 permits for certain species. Additionally, when external providers are slow to sequence
123 samples, it further delays the process. These limitations can hinder biomonitoring
124 projects and slow down the detection of potential invasive species. As a result, the
125 implementation of measures to mitigate their impacts may also be delayed
126 (Krehenwinkel et al. 2019; Egeter et al. 2022). Despite these limitations, the Illumina
127 MiSeq sequencing technology has been favored due to its lower error rate and well-
128 established bioinformatic procedures (Piper et al., 2019). Yet, Braukmann et al. (2019)
129 demonstrated similar performance in sequence quality and insect species recovery using
130 IonTorrent platforms (Ion Torrent PGM, and Ion Torrent S5), which are more affordable
131 and thus less often dependent on external providers and their sequencing delays.

132

133 In recent years, Oxford Nanopore Technologies® have released a very inexpensive
134 portable sequencing platform, the MinION. This small sequencer can be connected via
135 USB to a laptop to perform sequencing (Krehenwinkel et al. 2019) in the field and obtain
136 sequencing data in real time conditions. Indeed, the MinION for a metabarcoding
137 application offers the possibility of performing DNA sequencing of bulk samples directly
138 on site without the need for transport or relying on external sequencing providers. So
139 far, although the MinION does not seem suitable for the characterization of complex
140 communities, it is already suitable for the analysis of metabarcoding data when the
141 species diversity per sample is low and the target species are well represented in public
142 databases (Ho et al. 2020). In addition, recent developments in Nanopore technology and
143 base calling have reduced sequence error from 6% (Srivathsan et al 2021) down to less than
144 1% (Srivathsan et al 2024).

145

146 The primary objective of our study was to determine the most effective metabarcoding
147 approach for the biosurveillance of Cerambycid wood-boring beetles. To achieve this, we
148 compared the performance of three Next Generation Sequencing technologies: the
149 portable Nanopore sequencer MinION, the Illumina MiSeq and the Ion GeneStudio S5
150 (IonTorrent®). Our evaluation focused on their ability to detect invasive species in
151 different mock communities. Specifically, we assessed their accuracy in differentiating
152 between closely related cerambycid species and detecting low-abundance species in
153 mixed trap samples. Additionally, we analysed various metabarcoding primer pairs to
154 evaluate their accuracy in species identification. Finally, we emphasized the significance

155 of the field sampling protocol, particularly the trapping methods (dry versus
156 monopropylene glycol) in species detection.

157

158 **Materials and Methods**

159 *Taxa sampling*

160 Mock communities were constructed using 48 field-trapped specimens from different
161 countries in Europe (France, Greece, Portugal, Spain), China (Beijing and Zhejiang
162 Province) and USA (Michigan) (**Tab. 1**), as part of a worldwide trapping experiment
163 using multi-funnel traps baited with a generic attractant blend including eight
164 Cerambycid pheromones (see details of the blend composition and trapping methods in
165 Roques et al. 2023). Most of the specimens (36/48) were caught using α -cypermethrin
166 insecticide (Storanet®, BASF Pflanzenschutz Deutschland, Germany) in the trap basins,
167 of which the bottom had been replaced with a wire mesh to allow drainage and keep
168 specimens dry (hereafter called “dry” method). Other specimens (12/48) were captured
169 using a 50:50 ratio of monopropylene glycol (MPG) and water (hereafter called “wet”
170 method). Cerambycides collected from field-traps were stored in ethanol 95% and kept
171 at -20°C until further processing. Two individuals were captured by hand (“hand
172 collected” in Table 1) and pinned in collection boxes after capture. Date, country of
173 collection, type of trap and the 48 specimens used in mock communities are detailed in
174 Table 1.

175

176 *Mock community construction and DNA extraction*

177 Six mock communities with varying species composition were assembled as follows:

178

179 Test 1: Identifying closely related species.

180 To assess the efficiency of the different sequencing technologies and primers to
181 differentiate between sister species, bulks 1 and 2 were composed of congeneric species
182 (**Tab. 1**). Two legs from each individual (one specimen per species) were collected and
183 pooled to constitute the bulks. The whole set of legs was then ground using flame-
184 sterilized metal pestles to limit the risk of contamination. DNA from the ground material
185 was extracted using the DNeasy Blood and Tissue Kit (Qiagen) following the
186 manufacturer's instructions. Two additional legs were taken from the same specimens,
187 to assess the quantity and quality (A260/280 and A260/230 ratios) of DNA for each
188 specimen individually (**Fig. 1a**).

189

190 Test 2: Detecting low abundance species.

191 Bulks 3 and 4 were composed of six species represented by heterogeneous DNA
192 concentrations (**Tab. S1**) to assess the ability of the sequencing technologies and
193 primers to detect species present in a very low abundance. DNA of each individual (one
194 specimen per species) was previously extracted using two legs that were ground as
195 above and processed using the Qiagen DNeasy Blood and Tissue Kit. To construct bulks 3
196 and 4, individual DNA extracts were quantified using a fluorometer (Nanodrop™,
197 Thermo Fisher Scientific) and mixed together according to their concentration to achieve
198 the needed proportions of DNA for each individual (six individuals of different species
199 ranging from 41% to 3% for Bulk 3 and six individuals of different species ranging from
200 50% to 0.5% for bulk 4) (**Tab. 1, Fig. 1b**).

201

202

203 Test 3: Mimicking field trap content on species composition.

204 Bults 5 and 6 were built to reconstitute real trap contents by a collaborator involved in
 205 Cerambycidae trapping campaigns using multi-pheromonal traps (Roques et al., 2023).
 206 These bults include individuals from a number of species native to Europe usually found
 207 in the traps deployed there, with the addition of non-native species which have either
 208 already been introduced or at risk of being introduced in Europe (Bulk 5: 22 individuals
 209 of eight species, including one non-native (*Cordylomera spinicornis*); Bulk 6: 41
 210 individuals of 12 species including two non-native ones (*Xylotrechus stebbingi* and
 211 *Xylotrechus chinensis*) (**Tab. 1**). The DNA was extracted following the same protocol as
 212 for bults 1 and 2 where two legs were taken from each individual and ground together
 213 for DNA extraction (**Fig. 1a**).

214

Bulk	Species	Country of collection	Collection Year	Collection type
1	<i>Arhopalus fesus</i>	Portugal	2020	Cypermethrin insecticide (dry method)
1	<i>Arhopalus rusticus</i>	France	2021	Cypermethrin insecticide (dry method)
1	<i>Arhopalus syriacus</i>	Portugal	2019	Monopropylene glycol (wet method)
1	<i>Xylotrechus arvicola</i>	Portugal	2021	Cypermethrin insecticide (dry method)
1	<i>Xylotrechus chinensis</i>	Greece	2019	Cypermethrin insecticide (dry method)
1	<i>Xylotrechus stebbingi</i>	Greece	2019	Cypermethrin insecticide (dry method)
1	<i>Xylotrechus undulatus</i>	USA	2019	Monopropylene glycol (wet method)

2	<i>Monochamus galloprovincialis</i>	Portugal	2019	Monopropylene glycol (wet method)
2	<i>Monochamus sutor</i>	France	2019	Cypermethrin insecticide (dry method)
2	<i>Monochamus carolinensis</i>	USA	2019	Monopropylene glycol (wet method)
2	<i>Monochamus scutellatus</i>	USA	2019	Monopropylene glycol (wet method)
2	<i>Phymatodes amoenus</i>	USA	2019	Monopropylene glycol (wet method)
2	<i>Phymatodes testaceus</i>	USA	2019	Monopropylene glycol (wet method)
2	<i>Phymatodes varius</i>	USA	2019	Monopropylene glycol (wet method)
2	<i>Phymatodes aereus</i>	USA	2019	Monopropylene glycol (wet method)
2	<i>Phymatodes dimidiatus</i>	USA	2019	Monopropylene glycol (wet method)
3	<i>Pyrrhidium sanguineum</i>	France	2020	Cypermethrin insecticide (dry method)
3	<i>Xylotrechus stebbingi</i>	Spain	2021	Cypermethrin insecticide (dry method)
3	<i>Monochamus galloprovincialis</i>	Spain	2021	Cypermethrin insecticide (dry method)
3	<i>Xylotrechus chinensis</i>	Greece	2019	Cypermethrin insecticide (dry method)
3	<i>Chlorophorus glabromaculatus</i>	France	2020	Cypermethrin insecticide (dry method)
3	<i>Phymatodes testaceus</i>	France	2020	Cypermethrin insecticide (dry method)
4	<i>Arhopalus ferus</i>	France	2020	Cypermethrin insecticide (dry method)

4	<i>Monochamus sutor</i>	France	2019	Cypermethrin insecticide (dry method)
4	<i>Aegomorphus francottei</i>	France	2020	Cypermethrin insecticide (dry method)
4	<i>Monochamus galloprovincialis</i>	France	2018	Cypermethrin insecticide (dry method)
4	<i>Xylotrechus stebbingi</i>	Spain	2021	Cypermethrin insecticide (dry method)
4	<i>Xylotrechus chinensis</i>	Greece	2019	Cypermethrin insecticide (dry method)
5	<i>Pyrrhidium sanguineum</i>	France	2021	Cypermethrin insecticide (dry method)
5	<i>Batocera rubus</i>	China	2012	Hand collected
5	<i>Cerambyx scopolii</i>	France	2020	Cypermethrin insecticide (dry method)
5	<i>Cordylomera spinicornis</i>	France	2020	Cypermethrin insecticide (dry method)
5	<i>Leiopus femoratus</i>	France	2021	Cypermethrin insecticide (dry method)
5	<i>Leiopus nebulosus</i>	France	2020	Cypermethrin insecticide (dry method)
5	<i>Pachyta bicuneata</i>	China	1987	Hand collected
5	<i>Stictoleptura cordigera</i>	France	2021	Cypermethrin insecticide (dry method)
6	<i>Arhopalus rusticus</i>	France	2020	Cypermethrin insecticide (dry method)
6	<i>Xylotrechus chinensis</i>	Greece	2019	Cypermethrin insecticide (dry method)
6	<i>Plagionotus detritus</i>	France	2020	Cypermethrin insecticide (dry method)
6	<i>Plagionotus arcuatus</i>	France	2020	Cypermethrin insecticide (dry method)

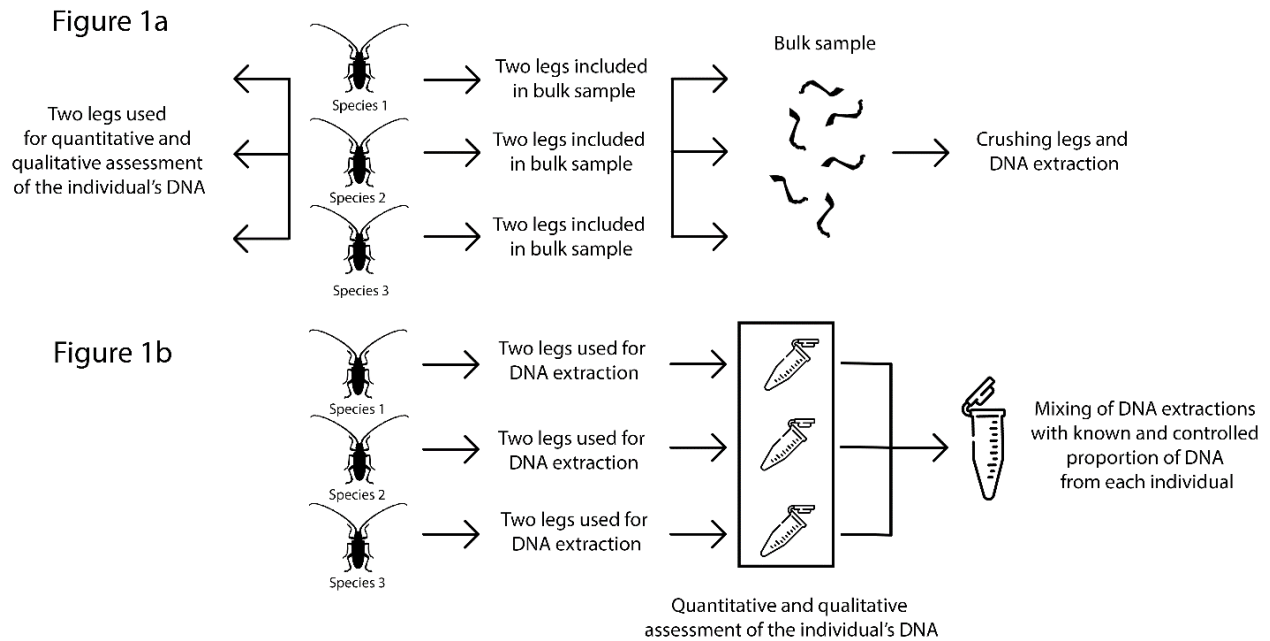
6	<i>Xylotrechus stebbingi</i>	France	2020	Cypermethrin insecticide (dry method)
6	<i>Arhopalus syriacus</i>	France	2020	Cypermethrin insecticide (dry method)
6	<i>Arhopalus fesus</i>	France	2020	Cypermethrin insecticide (dry method)
6	<i>Xylotrechus colonus</i>	USA	2019	Monopropylene glycol (wet method)
6	<i>Chlorophorus ruficornis</i>	France	2021	Cypermethrin insecticide (dry method)
6	<i>Phymatodes testaceus</i>	France	2021	Cypermethrin insecticide (dry method)
6	<i>Prionus coriarius</i>	France	2010	Cypermethrin insecticide (dry method)
6	<i>Phymatodes amoenus</i>	USA	2019	Monopropylene glycol (wet method)

215

216 Table 1: Species, origin, date and condition of capture of the specimens used in the 6
 217 bulks used. Species names in bold correspond to exotic species. We consider specimens
 218 that have been captured on a different continent than their place of origin as exotic.

219

220



221

222

223 Figure 1: Overview of the DNA extraction protocol for tests 1 (identifying closely related
 224 species) and 3 (mimicking field trap content on species composition) (Fig. 1a) and for
 225 test 2 (detecting low abundance species) (Fig. 1b).

226

227

228 *PCR amplification*

229 All bulks samples were amplified with two pairs of primers internal to the commonly

230 used barcode fragment : BF3/BR2 (called hereafter "B") (CCHGAYATRGCHTTYCCHCG /

231 TCDGGRTGNCCRAARAAYCA (Elbrecht and Leese 2017; Elbrecht et al. 2019), which

232 generates a 458 bp was used for all the technologies; and fwhF2/fwhR2n (called

233 hereafter "F") (GGDACWGGWTGAACWGTWTAYCCHCC /

234 GTRATWGCHCCDGCTARWACWGG), which generates a shorter 254 bp amplicon (Vamos

235 et al. 2017) was used for Illumina and MinION technologies only. Each PCR comprised

236 15.3 µl H₂O, 2.5 µl 10X PCR buffer, 2.5 µl dNTPs [1mM], 1 µl of each primer [0.4mM], 0.2

237 µl Dream Taq (Thermo Fisher Scientific), 0.5 µl Betaine [100mM] and 2 µl DNA for a total

238 of 25 µl per reaction. For both primer pairs, PCR was performed using the same

239 program: 95°C for 5 min, 29 cycles of 95°C for 30 s, 48°C for 30 s, and 72°C for 50 s and

240 72°C for 5 min (Elbrecht et al. 2019). PCR products were then run on a 2% agarose gel
 241 stained with ethidium bromide and visualised by UV transilluminator. The PCR products
 242 were then purified with the NucleoFast 96 PCR plate purification kit (Macherey-Nagel).
 243 Three PCR replicates were performed for the six bulks when using the MinION technology.

244

245 *Illumina® library preparation*

246 A second ligation PCR was performed on the products of the first PCR to add Illumina®
 247 tags and adapters, prepared by ligating Nextera XT indices through an eight cycle PCR
 248 (with a modified PCR protocol). The second PCR was carried out with the same condi-
 249 tions as for the initial PCR. Reactions (25 µl) contained the following: 5 µl of template
 250 DNA (purified products from the first PCR), 1 µl of each primer [10 µM], 5 µl of 5X GoTaq
 251 (Promega) reaction buffer, 1µl of MgCl₂ [25mM], 1 µl of BSA [1 mg/ml], 0.5 µl of dNTPs
 252 [5mM], 0.125 µl of GoTaq G2 Polymerase (Promega) and 10.375 µl of molecular-grade
 253 water to reach 25 µl. The PCR conditions were the same as for the first PCR, with eight
 254 cycles. The products of the second PCR were verified on a 2% agarose gel. PCR products
 255 were then equimolarly pooled into two different pools (one pool per primer pair used)
 256 and purified using the GeneJET Gel Extraction kit from an agarose gel, following manu-
 257 facturer's instructions. This library was sequenced in Illumina MiSeq using V3 chemistry
 258 (300 × 300 bp, 600 cycles) in the Sequencing Center within the Biozentrum of the Lud-
 259 wig-Maximilian University in Munich (Germany).

260

261 *MinION library preparation*

262 Libraries were prepared according to the Oxford Nanopore Technologies ® protocol:
 263 "PCR barcoding (96) amplicons (SQK-LSK110) (version:

264 PBAC96_9114_v110_revF_10Nov2020)" with the following specifications. After the first
265 PCR describe above, the Nanopore PCR barcoding expansion Pack 1-96 (EXP-PBC096)
266 was used to perform the second PCR to incorporate the Oxford Nanopore Technologies
267 ® barcode sequences on the amplicons generated in the first PCR.

268

269 Reactions (50 µl) contained the following: 2 µl of template DNA (purified products from
270 the first PCR), 0.5 µl of each primer [10 µM], 10 µl of 5X GoTaq (Promega) reaction buff-
271 er, 2 µl of MgCl₂ [25 mM], 2 µl of BSA [1 mg/ml], 2 µl of Q solution, 1 µl of dNTPs [5 mM],
272 0.3 µl of GoTaq G2 Polymerase (Promega) and 29.7 µl of molecular-grade water to reach
273 25 µL. The thermocycling conditions followed the manufacturer recommendations: 95°C
274 for 3 min, followed by 15 cycles of 95°C for 15 s, 62°C for 15 s, and 65°C for 30 s and 65°C
275 for 7 min.

276

277 Final PCR products were then quantified using Qubit and equimolarly pooled before
278 being purified with Agencourt AMPure XP beads (Beckam Coutler). The final pool was
279 then sequenced on the MinION sequencer (Mk1c; Oxford Nanopore Technologies ®, UK)
280 using a R10.3 flowcell (MIN111) with 1331 pores available and the LSK110 ligation
281 sequencing kit. The two replicates of bulk 6 using the MinION technology were of
282 insufficient quality (Nanodrop) and were therefore removed from the analysis.

283

284 *IonTorrent® library preparation*

285 For the production of the libraries, we started with 5 ng of DNA extract (Qubit
286 measurement). The Nextflex Cellfree DNaseq kit (PerkinElmer) was used for the
287 process. The quality of the libraries was assessed using Qubit (for quantification) and

288 Bioanalyzer (using the HighSensitivity kit from Agilent, for size verification). After
289 quality control, each library was amplified by emulsion PCR on the Ion One Touch 2
290 instrument, with a concentration of 15 pg/ μ l. Subsequently, the libraries were
291 sequenced on an Ion GeneStudio S5 system using a single-end sequencing protocol with
292 a 300 bp read length. Sequencing was performed on an Ion 520 Chip by the GeT-
293 BioPuces platform (Toulouse, France).

294

295 *False positive detections*

296 False positive detections are considered as the detection of a species within a bulk that
297 was not initially present when the bulks were constructed. In order to estimate the
298 representativeness of false positives within true positives in the bulks, the total number
299 of reads assigned to false positive OTUs was reconciled and compared to the number of
300 reads assigned to non-false positive detections. The number of false positives detected
301 according to the different tested combinations are indicated in **Tab. S2**.

302

303 *Reference Barcode Dataset*

304 A dataset was built using all the public sequences of all Cerambycidae species available
305 in BOLD systems v4 (Ratnasingham and Hebert 2007). It was then verified whether all
306 33 species present in the bulk samples were represented by at least one sequence in the
307 database. Three species not previously included in the database were barcoded through
308 Sanger sequencing on an ABI 3500 genetic analyzer (Applied Biosystems) using the Big-
309 Dye Terminator V3.1 sequencing kit (Applied Biosystems) and BF3/BR2 primer pair.
310 They were subsequently added to our local database to ensure that they were

311 represented by at least one barcode sequence. The number of sequences in the database
312 for each species is shown in **Tab. S1**.

313

314 The final reference data set is available from BOLD in the dataset DS-MINION
315 (dx.doi.org/10.5883/DS-MINION) and includes once barcode per species together with
316 the three newly generated barcodes. Lab Sheet from the DS-MinION database is shown
317 in **Tab. S3**.

318

319 *Illumina® data processing*

320 The raw data was analysed using the *FROGS v4.0.1* pipeline, a standardized pipeline
321 containing a set of tools that are used to process amplicon reads that have been
322 produced from Illumina® sequencing (Escudié et al. 2018; Henrie et al. 2022). First,
323 amplicons with a size between 408 and 508 for the BF3/BR2 primer pair and 204 and
324 304 for the fwhF2/fwhR2n primer pair were retained using Pre-process tool. Sequence
325 clustering was then performed using the SWARM algorithm (Mahé et al. 2014) with a
326 maximum sequence difference set at $d=1$ (--distance 1 parameter), as recommended by
327 SWARM. Chimeric sequences were then removed with Remove chimera tool. Sequences
328 were aligned to the same database used for the MinION and IonTorrent® data analysis.
329 In order to remove all spurious detections, OTU detections with less than 10 reads were
330 removed. In barcoding and metabarcoding studies of insects, the sequence similarity
331 level for OTU identification usually ranges from 95% to 99% (e.g., Gibson et al., 2014;
332 Zenker et al., 2016). We calculated the best threshold value for our dataset by applying
333 the function *localMinima* from R package *spider v1.5.0* (Brown et al., 2012). Based on this

334 analysis, we used a threshold of 98% to assign OTUs to species level. The resulting OTU
 335 tables for Illumina F and Illumina B are provided in **Tab. S4** and **Tab. S5**, respectively.

336

337 *MinION and IonTorrent® data processing*

338 Bioinformatics analyses were performed on the Genotoul Bioinformatics Platform
 339 (INRAE, Toulouse, France). Basecalling and demultiplexing were performed for MinION
 340 data using Guppy v6.1.7; ONT; high accuracy base calling mode; parameters: -c
 341 dna_r10.3_450bps_hac.cfg --min_qscore 5 --trim_barcodes. Then, for MinION and
 342 IonTorrent® data, we used the *msi* data processing pipeline v0.3.6 (Egeter et al. 2022) to
 343 reduce the error rate of the reads by polishing them after the basecalling step. Reads
 344 smaller than 40bp were removed with cutadapt v4.0 (Martin 2011). The size range was
 345 set between 408bp and 508bp for BF3/BR2 and between 204bp and 304bp for
 346 fwhF2/fwhR2n. The clustering step was carried out with *ISONCLUST* v0.0.6.1 (Sahlin and
 347 Medvedev 2020; with parameters: --mapped_threshold 0.825 and --aligned_threshold
 348 0.55) and a consensus sequence per cluster was generated using RACON v1.5.0 (Vaser et
 349 al. 2017). The polished reads were then clustered at 97% sequence identity with *CD-HIT*
 350 v4.8.1 (Fu et al. 2012) and a representative sequence from each cluster (centroid) was
 351 selected. The polished reads were then aligned to the local database with BLAST
 352 (BLASTn algorithm). The following parameters were used: -word_size 11 -perc_identity
 353 95 -qcov_hsp_perc 98 -gapopen 0 -gapextend 2 -reward 1 -penalty 1 -max_target_seqs
 354 100. Similarly, to the Illumina® data processing, OTU detections with less than 10 reads
 355 were removed. Finally, a taxonomic assignment was performed for each query using a
 356 Lowest Common Ancestor (LCA) approach with the bioinformatics package metabinkit
 357 (Chain et al. 2016; Egeter et al. 2018; Kitson et al. 2019) with the following thresholds:

358 98% at species level, 97% at genus level, 95% at family level (Alberdi et al. 2018; Egeter
359 et al. 2022). The resulting OTU tables for MinION B, MinION F and IonTorrent B are
360 provided in Tab. S6, Tab. S7 and Tab. S8 respectively.

361

362 *Statistical analysis*

363 A two-sample test of proportions was used to compare and assess the significance of the
364 proportion of reads assigned to the species levels for MinION, Illumina, and IonTorrent
365 technologies using the "Social Science Statistics" website
366 (<https://www.socscistatistics.com/tests/anova/default2.aspx>). The proportion of reads
367 assigned to different taxonomic levels was calculated by summing the total reads from
368 different bulk samples for each condition. To determine if the number of false positives
369 was significantly different among the three technologies and the two primer pairs, we
370 calculated the detection mean for each bulk under different conditions. We then
371 performed an ANOVA test followed by a Tukey HSD test using the "Social Science
372 Statistics" website. The Wilcoxon test, Exact Fisher's test and standard deviation were
373 calculated in R v4.3.2 (<https://www.R-project.org/>).

374

375 **Results**

376 A total of 1,248,95 reads were sequenced with the MinION Nanopore® technology using
377 the F primer pair, with an average of 78,037 (SD=28,415) reads per sample. After quality
378 filtering, and removal of reads of incorrect size or insufficient quality, 1,113,844 (89.2%)
379 reads were retained, with an average of 69,615 (SD=25,508) reads per bulk. For the B
380 primer pair, a total of 1,132,604 reads were sequenced, with an average of 62,922

381 (SD=17,442) reads per sample. After quality filtering, a total of 948,832 (83.8%) reads
 382 were retained, with an average of 52,712 (SD=14,512) reads per bulk (**Tab. 2**).

383

384 The Illumina® sequencing produced a total of 1,549,894 reads using the B primer pair,
 385 with an average of 258,316 (SD=39,365) reads per bulk. After quality filtering, 1,025,637
 386 (66.2%) reads were retained, with an average of 170,940 (SD=69,961) reads per bulk.
 387 For the F primer pair, a total of 2,383,028 reads were sequenced, with an average of
 388 397,171 (SD=84,482) reads per bulk. After quality filtering, 1,686,058 (73.3%) reads
 389 were retained, with an average of 281,010 (SD=112,512) reads per bulk (**Tab. 2**).

390

391 Regarding the IonTorrent® technology, 838,489 reads were sequenced, with an average
 392 of 139,748 (SD=17,086) reads per bulks with the B primer pair. After the quality
 393 filtering, 280,695 (33.5%) reads remains with an average of 46,782 (SD=5,025) reads
 394 per bulks (**Tab. 2**).

395

Technology	Primer pair	n_raw_reads	n_reads_post_filtering
MinION	B	1,132,604	948,832
	F	1,248,95	1,113,844
Illumina	B	1,549,894	1,025,637
	F	2,383,028	1,686,058
IonTorrent	B	838,489	280,695

396

397 Table 2: Number of raw reads obtained after sequencing and after pre-process steps
 398 according to sequencing technologies and primer pairs used.

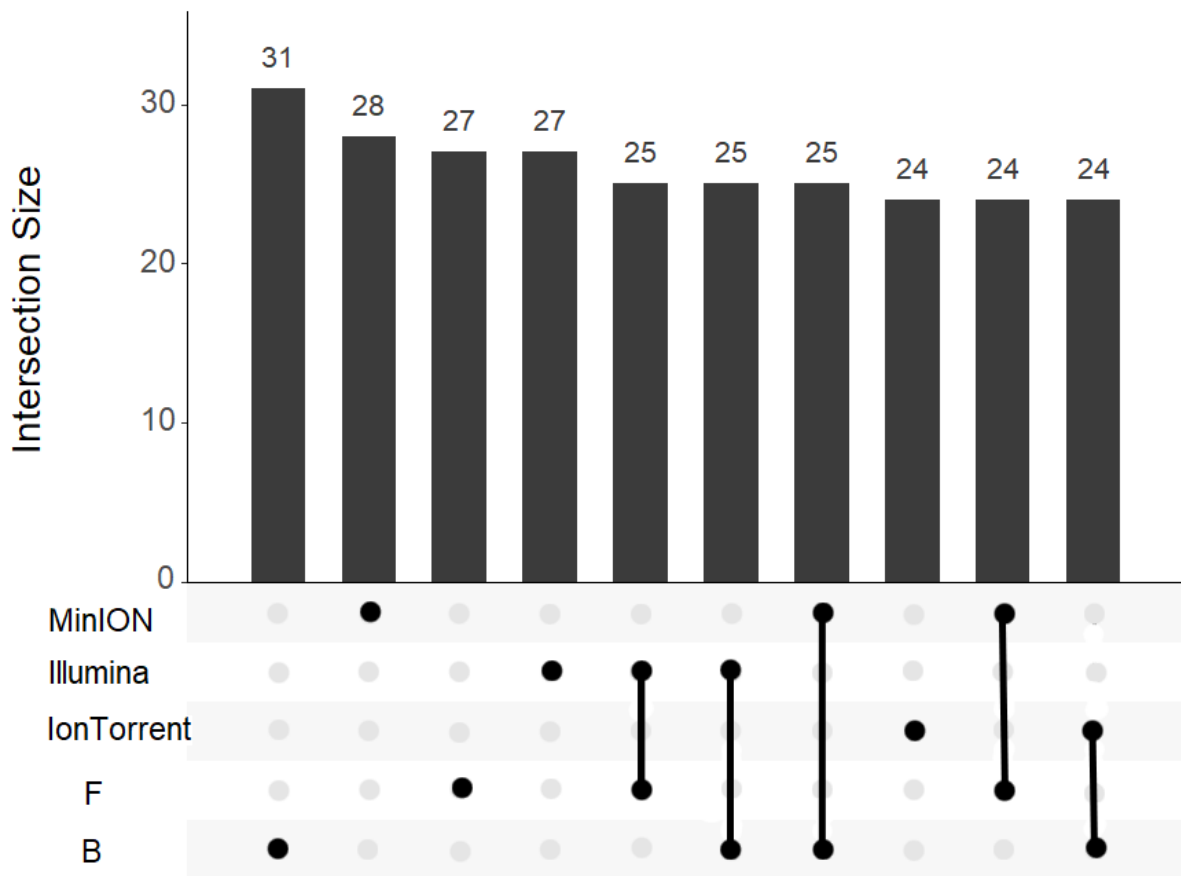
399

400 *Benchmarking of sequencing technologies*

401 The MinION technology accurately identified 28 out of 48 specimens at the species level,
 402 Illumina® technology allowed specific identification of 27 specimens and IonTorrent®

403 identified 24 specimens. The primer pair F enabled the specific identification of 27
 404 specimens at species level while the primer pair B enabled the identification of 31
 405 specimens at species level. Illumina® F, Illumina® B and MinION B allowed for 25
 406 species-level identifications across all bulks and 24 for MinION F and for IonTorrent® B.
 407 This difference was not significant (Fisher's Exact Test, $p = 1.00$) (**Fig. 2**).

408



409

410

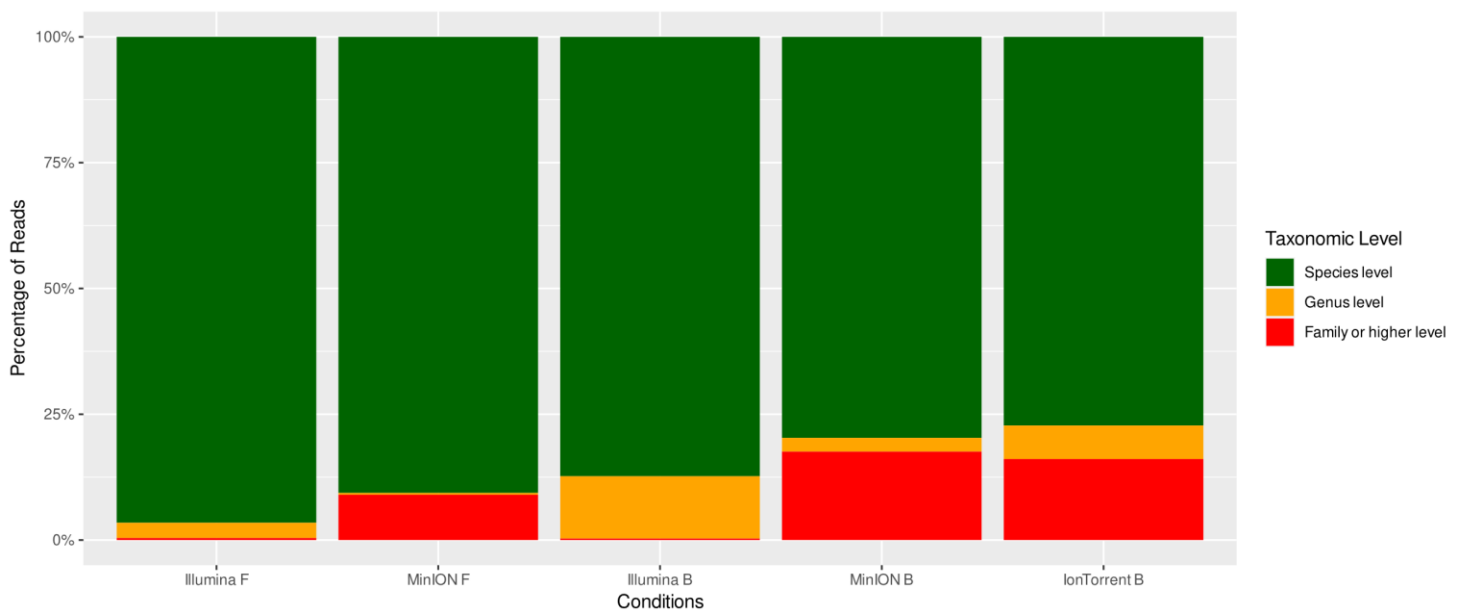
411 Figure 2: Upset plot showing the number of individuals detected at species level
 412 according to the three technologies (Illumina, MinION and IonTorrent), primer pairs
 413 (F=fwhF2/fwhR2n [254bp] and B=BF3/BR2 [458bp]) and technology-primer pair
 414 combinations tested.

415

416

417 The number of reads assigned at the species level was significantly higher with
 418 Illumina® technology ($p.value < 0.00001$) compared to MinION. Nearly 97% of reads
 419 were assigned at the species level for the Illumina® F combination compared to 90% for

420 the MinION F combination (p.value < 0.0001). As for primer pair BF3/BR2, over 87.3%
 421 of reads were assigned at the species level for Illumina®, followed by over 79.7% for
 422 MinION technology and over 77.2% for IonTorrent® technology (**Fig. 3**). The primer
 423 pair fwhF2/fwhR2n resulted in a significantly higher percentage of reads assigned at the
 424 species level (93.6%) (considering both Illumina® and MinION technologies) compared
 425 to couple of primers B (81.4%) (considering all three technologies) (p.value < 0.00001).
 426 Summary table of the number of reads (**Tab. S8a**) or percentage (**Tab. S8b**) assigned to
 427 the taxonomic levels of species, genus and family or higher or all conditions tested are
 428 shown in **Tab. S9a** and **Tab. S9b**

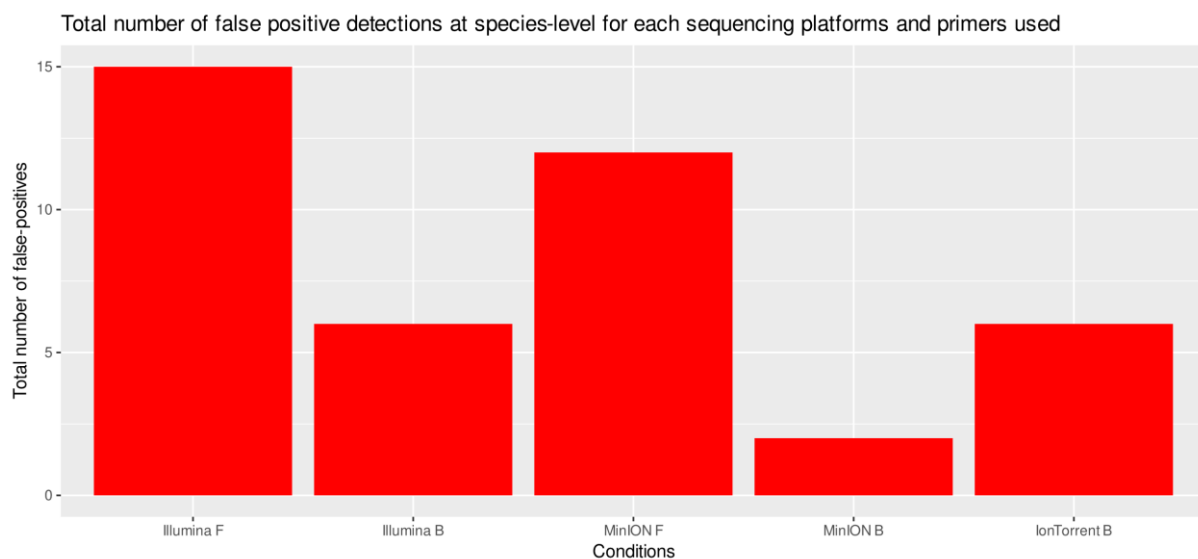


429
 430 Figure 3: Proportion of reads assigned to each taxonomic level for each combination of
 431 sequencing technology and pair of primers (F: fwhF2/fwhR2n; B: BF3/BR2).

432
 433 False positive detections (i.e. a species detected within a bulk that is not part of the
 434 bulk's initial composition) were observed whatever the combination of primers and
 435 technology (**Fig. 4**). Hence, an average of 13.5 false positives OTU were recorded for the
 436 primer pair fwhF2/fwhR2n, compared to an average of four false positives OTU when

437 using the primer pair BF3/BR2, the difference being significant here (p.value = 0.00194).
 438 According to the technology used, but regardless of the primers, an average of ten, seven
 439 and six false positives were recorded for Illumina®, MinION and IonTorrent®
 440 technologies respectively. There are no significant differences among the three
 441 sequencing technologies in terms of false positives.

442



443 Figure 4: Number of false positive detections at species-level for each sequencing
 444 platforms and primers used (F=fwhF2/fwhR2n [254bp] and B=BF3/BR2 [458bp]).

445

446

447 *Mock community analysis*

448 In total, 33 out of 48 individuals (68.8%) were detected at the species level by at least
 449 one experimental condition (**Fig. 5**).

450

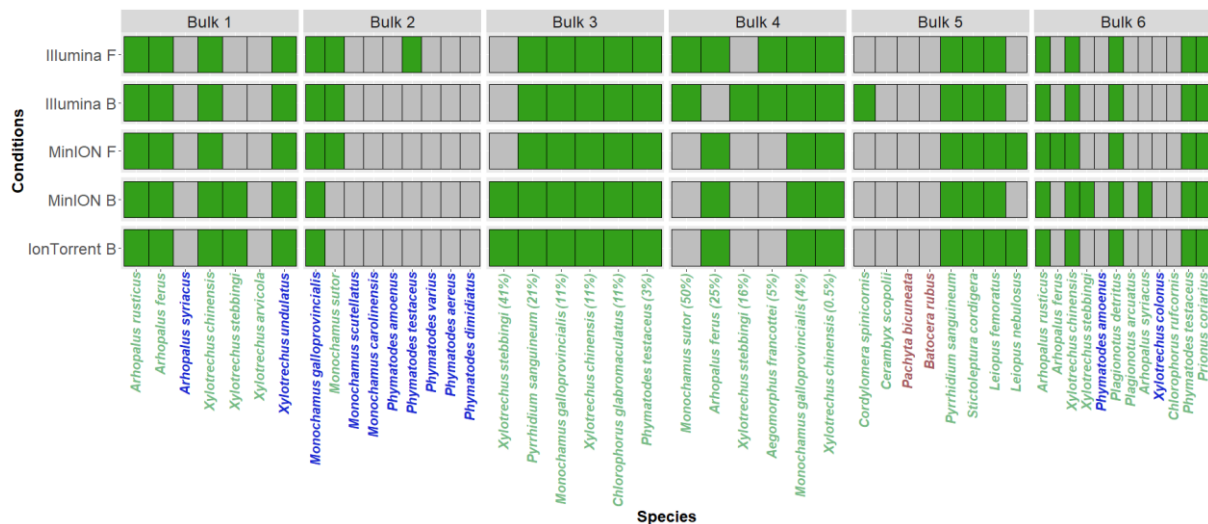
451 Bulks 1 and 2 were assembled to compare the detection rates of closely related species
 452 under different sequencing and primers conditions. Illumina® detected seven species
 453 out of 16 (43.75%), MinION also detected seven out of 16 (43.75%), and IonTorrent®
 454 detected six species out of 16 (37.5%). No significant differences were observed among
 455 the different methods used (Kruskal-Wallis chi-squared = 2, df=2, p-value=0.3679).

456 Metabarcoding of bulks 3 and 4 aimed at comparing the ability of different sequencing
 457 technologies to detect low abundance species in the traps. All sequencing
 458 technology/primer combinations allowed for the detection of minor species:
 459 *Phymatodes testaceus* with a presence of 3% in bulk 3 (relative amount of DNA in the
 460 mock community) and *Xylotrechus chinensis* with a percentage of 0.5% in bulk 4.
 461 However, some species (although not in minority in the bulks) were not detected in one
 462 or several tests (**Fig. 5**). In total, Illumina® was able to detect a higher number of
 463 individuals (11/12 individuals detected) compared to MinION (9/12) and IonTorrent®
 464 (9/12).

465

466 Regarding bulks mimicking the species composition in a field trap, MinION performed
 467 better to detect and identify specimens at species level in Bulk 6 (detecting 8/12 species
 468 (66.7%)) compared to Illumina® and IonTorrent® technologies (5/12 species (41.7%)),
 469 whereas the same number of species was detected for Bulk 5 (4/6 (66.7%)) whatever
 470 the technology used. Nevertheless, in bulk 5, the non-native species, *Cordylomera*
 471 *spinicornis* was detected only by Illumina B. For bulk 6, the non-native species
 472 *Xylotrechus chinensis* was detected by all three technologies and *Xylotrechus stebbingi* by
 473 MinION B only.

474



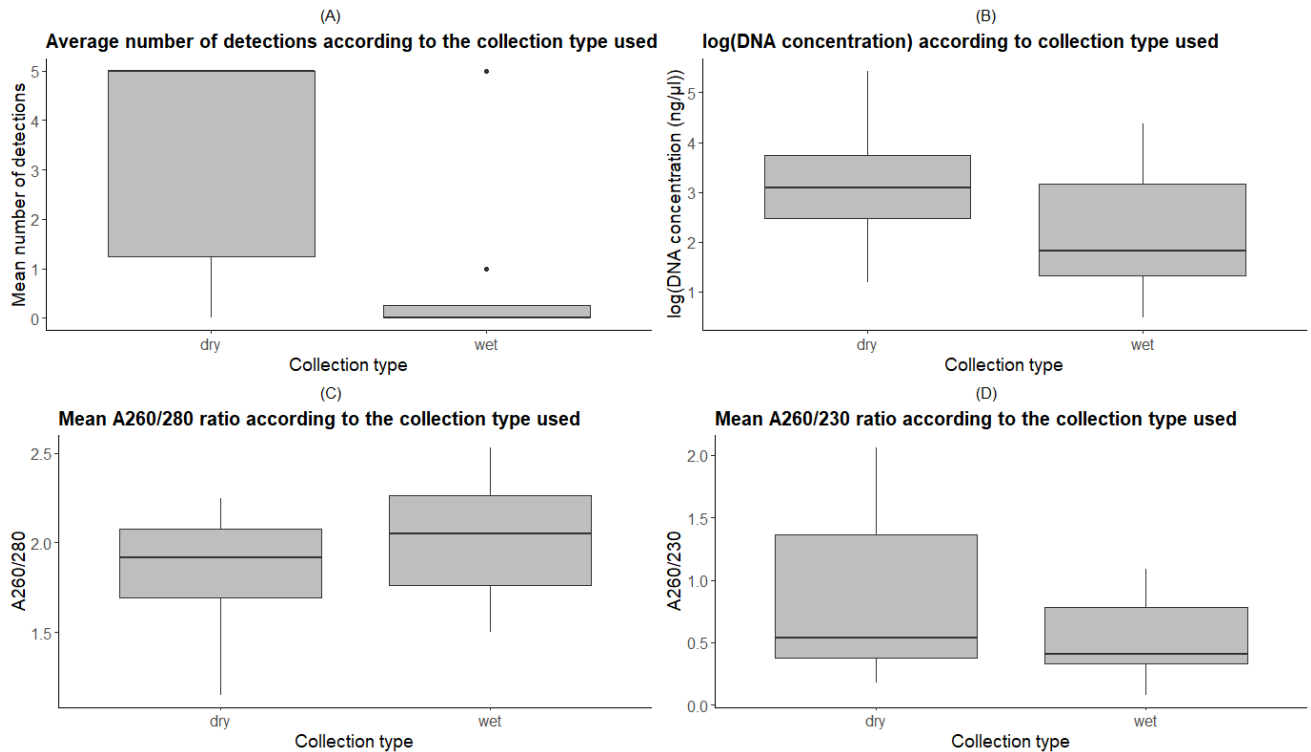
475
 476 Figure 5: Heatmap comparing the identification of individuals present in bulk samples at
 477 the species level (green square) or the absence of detection at the species level (grey
 478 square) according to the sequencing technologies and primer pairs used (F=fwhF2/FwhR2n; B=BF3/BR2). Species names written in blue were collected using
 479 the wet method, those in green were collected using the dry method, and those in dark
 480 red were hand-captured.
 481
 482

483 *Impact of capture and storage conditions on individual detection*

484 Our results demonstrate significant differences in the mean number of detections
 485 between samples that were collected using the “dry” method (α -cypermethrin
 486 insecticide) and the “wet” method (water-diluted propylene glycol) (Wilcoxon rank-sum
 487 test, $W = 74.5$, $p.value = 0.0006342$) (Fig. 6A). Indeed, 75% (9/12) of specimens
 488 collected using the “wet” MPG trapping procedure, were not detected by any of the
 489 sequencing technologies. Conversely, 88.2% (30/34) of those collected using the “dry”
 490 trapping procedure (based-on α -cypermethrin insecticide) were detected at least once
 491 across all technologies.

492
 493 Individuals captured using the "dry" method had higher DNA concentration (39 ng/ μ l on
 494 average, (SD=52.79)) than MPG trapped specimens (18.6 ng/ μ l on average, (SD=21.80))
 495 (Wilcoxon rank-sum test, $W = 123.5$, $p.value = 0.04533$) (Fig. 6B). The average

496 A260/280 ratio was 1.9 for the “dry” method and 2 for MPG method (**Fig. 6C**). However,
 497 the average A260/230 ratio of specimens trapped with the “dry” method (0.8) was
 498 higher than that of specimens captured with MPG (0.5) (Wilcoxon rank-sum test, $W =$
 499 146, p .value = 0.1502) (**Fig. 6D**).
 500



501
 502 Figure 6: Boxplots representing (A) the average number of detections according to the
 503 type of preservation used, (B) the natural logarithm scale (base e) of the average DNA
 504 concentration according to the type of preservation used, (C) the A260/280 quality ratio
 505 according to the type of preservation used, and (D) the A260/230 quality ratio according
 506 to the type of preservation used. The black dots represent the outliers values (values
 507 outside the whiskers). Bold line represents the average value, outlines of the boxes
 508 represent the first and third quartiles and the whiskers represent the range of the values
 509 outside the quartiles.
 510

511 **Discussion**

512 Rapid and precise detection of exotic insects is crucial to prevent the ecological and
 513 economic damage they can cause by invading new environments and disrupting local
 514 ecosystems.

515

516 *Benchmarking of sequencing technologies*

517 A slightly higher number of individuals were detected and identified to species using
518 MinION (28 specimens) compared to Illumina® (27 specimens) or IonTorrent® (24
519 specimens), although this difference is not significant. However, this result shows similar
520 detection rates between the MinION and Illumina technology, demonstrating that the
521 sequencing error rates long attributed to the MinION did not impact detection rates,
522 while allowing for the elimination of the long delays often required when sequencing is
523 performed on other sequencing technologies (Piper et al., 2019). It must be considered
524 that we worked on a single pair of primers (BF3/BR2) with the IonTorrent® technology,
525 which may have reduced the number of identifications. More specifically, our results
526 showed that the choice of primer pairs, and the length of the amplicon generated, lead to
527 contrasted results regarded to taxonomic assignment. For example, only BF3/BR2
528 allowed the species-level identification of the invasive species *Xylotrechus stebbingi*. This
529 difference may be due to the longer amplicon generated by this primer pair, which has
530 more informative nucleotide sites to provide a reliable taxonomic assignment. By
531 contrast, fwhF2/fwhR2n generated a significant higher number of false positives than
532 BF3/BR2 (**Fig. 4**). This may be because the amplicon generated by fwhF2/fwhR2n is
533 smaller in size compared to BF3/BR2. As a result, any loss of genetic information is more
534 likely to result in misidentification or false positives (Meusnier et al., 2008).

535

536 Regardless of the number of identified species, the Illumina® technology produced a
537 higher percentage of reads allowing species-level identification compared to MinION or
538 IonTorrent®. The detection of specimens at higher taxonomic level (genus or family),

539 can be explained by sequencing errors that produce reads with less than 98% identity to
 540 the reference database. These results confirm that Illumina has a lower sequencing error
 541 rate than Oxford Nanopore's MinION sequencer (Piper et al; 2019), although this did not
 542 impact the number of individuals identified to the species level.

543

544 The three technologies showed similar efficiency in detecting and identifying closely
 545 related species. Moreover, the results show that all three sequencing technologies
 546 (regardless of the associated primer pairs) enabled the detection and identification of
 547 species whose DNA represented a very low percentage in the mock community (**Fig. 5**).
 548 This high resolution would allow for the detection of exotic species that are poorly
 549 represented in traps, which might otherwise go unnoticed. Thus, all three technologies
 550 appear suitable for detecting and identifying species present in low numbers in field
 551 traps, enabling effective monitoring.

552

553 *Impacts of capture and storage conditions on DNA conservation*

554 Both the conditions of capture (wet versus dry methods) and storage (ie time lag
 555 between collection and lab processing) have an impact on DNA concentration and
 556 quality and subsequently on the rate of species detection (Piper et al., 2019). Thus, the
 557 number of species detected is highly variable between bulks 1 and 2, which can be
 558 explained by the capture methods used : 'dry,' where individuals were captured without
 559 preservative fluid (as is the case for the majority of detected individuals comprising bulk 1),
 560 and 'wet,' where individuals were preserved in 50% MPG until trap retrieval (as is the
 561 case for the majority of undetected individuals comprising bulk 2). For instance, the species
 562 *Phymatodes testaceus* was always detected (10 out of 10 assays) when dry specimens

563 were present, even in low concentrations (3% in Bulk 3). On the other hand, wet
 564 specimens of *Phymatodes testaceus* were rarely detected (one detection out of five
 565 assays). Individuals captured using MPG method had lower DNA concentration and
 566 presented significantly much lower detection rates compared to individuals captured
 567 using the “dry” method. Ballare *et al.* (2019) also found that insects collected in
 568 propylene glycol traps produced lower quality ddRADseq assemblages than specimens
 569 collected by net sampling and directly transferred into 100% EtOH or by passive
 570 trapping followed by 100% EtOH storage before pinning. In contrast to this, Ferro and
 571 Park (2013) found that propylene glycol is an effective DNA preservative for molecular
 572 marker-based studies on Coleoptera species. However, in their study, insects were first
 573 killed and preserved in 100% ethanol before being stored in glycol. While, in our study
 574 insects were killed directly in propylene glycol. The use of 100% ethanol as the initial
 575 killing agent may lead to better initial preservation of specimens compared to direct
 576 exposure to propylene glycol.

577

578 *False positives, negatives and unmatched OTUs*

579 Despite the precautions taken, several false positives were detected in all tested
 580 conditions. The number of false positives was significantly higher with the primer pair
 581 fwhF2/fwhR2n, which generates a smaller size amplicon compared to BF3/BR2. Even
 582 though Illumina technology is known to have a lower sequencing error rate compared to
 583 MinION (Silvestre-Ryan and Holmes 2021), our study found 10 false positives generated
 584 by Illumina, while MinION produced seven false positives and IonTorrent produced six.
 585 The sensitivity of HTS technologies allows for the detection of very small amounts of
 586 DNA, thus detecting even the slightest cross contamination between samples (Liu et al.

587 2020). These DNA contaminations may have occurred during sample collection in the
588 field or in the laboratory through cross-contamination between samples from the same
589 study.

590

591 The false negative detections for some individuals may primarily be explained by the
592 highly heterogeneous DNA quality of the different sequenced individuals (**Tab. S1**). In
593 fact, DNA quality can be impacted by numerous mainly abiotic factors (pH, UV radiation,
594 temperature), degrading DNA quality in a matter of days/weeks (Strickler et al. 2015;
595 Collins et al. 2018; Harrison et al. 2019). During field trapping using stationary traps,
596 captured insects are sometimes exposed to such conditions (high temperatures in trap
597 containers when exposed to the sun in summer, high humidity in the container during
598 heavy rains, etc...), which can greatly accelerate the speed of DNA degradation in
599 captured individuals. Such degraded DNA is more difficult to amplify, thus generating
600 false negatives, especially when attempting to detect insects in low abundance within a
601 trap, such as an invasive species in the process of establishing (Preston et al. 2022).
602 Another possible cause for the high number of false negatives is the bias induced by PCR,
603 such as uneven amplification of the DNA of the different individuals present in one
604 sample (Preston et al. 2022). To avoid potential bias arising from identification errors or
605 missing species in the references databases, we decided to work on a local and curated
606 BLAST database. However, when target species are partially unknown, as is the case in
607 field conditions, analyses must rely on public reference databases. Yet, out of the 35,000
608 known species of Cerambycidae to date, only 2,926 species (8.4%) are recorded in BOLD
609 with a barcode fragment (as of November 16th, 2023). Furthermore, databases can
610 contain errors such as misassignment of an DNA sequence to a wrong species due to

611 morphological identification errors. This was precisely the error encountered for the
 612 species *Monochamus sutor* which was genetically identified as *Monochamus sartor* (**Tab.**
 613 **S7**) or the species *Leiopus nebulosus* who has been genetically identified as *Leiopus linnei*
 614 (**Tab. S7**) using our local BOLD database.

615

616 One also needs to pay attention to synonymy whereby a species appears in the database
 617 under multiple names. We encountered this problem in our analysis with *Arhopalus ferus*
 618 (Bulks 1, 4, and 6) which was detected but under the name of *Arhopalus tristis* (**Tab.**
 619 **S10**). Finally, mitochondrial paralogues such as NUMTs (non-functional copies of
 620 mitochondrial genes transported into the nuclear genome) present in databases can also
 621 bias results, making it impossible to identify specimens correctly at the species level
 622 (Bensasson et al. 2001). NUMTs are numerous in many organisms, including some
 623 cerambycids such as *Monochamus galloprovincialis* (Koutroumpa et al. 2009; Haran et al.
 624 2015).

625

626 *Biases*

627 Based on the results obtained, it appears that the main biases observed in
 628 metabarcoding analyses of trap contents stem from the degradation of DNA from
 629 individuals, which generates false negatives. We the recommend to favour a "dry" rather
 630 than a "wet" trapping method, especially the MPG method and to plan for the collection,
 631 transportation, and processing of captured individuals as soon as possible after capture.
 632 This includes checking the traps as frequently as possible (at least once a week), thus
 633 avoiding excessively long exposure of the individuals to unfavourable environmental
 634 conditions. Once individuals are brought back to the laboratory and if DNA cannot be

635 extracted straight away, it is important to limit any further degradation by keeping
636 samples at -20°C and in 95% ethanol. On the other hand, DNA extractions should be
637 stored in the preservation buffer provided with the extraction kits or in molecular-grade
638 water and kept at -20°C (Preston et al. 2022). We also recommend limiting the use of
639 primer pairs that generate short amplicons, which can favour the amplification of non-
640 target taxa, NUMTs and lead to identification errors. The quality and completeness of the
641 databases is also a very important bias factor. To limit this bias, Egeter et al., 2019
642 recommended to restrict the database used to targeted species in order to minimize the
643 risk of false positives due to contamination. Limited taxonomic and geographical
644 coverage of sequence databases is a huge limitation in metabarcoding studies. For
645 example, Dopheide et al. (2019) found no representative sequence in the GenBank
646 database for more than 900 invertebrate OTUs in their study when analysing the
647 community of soil arthropods from a native forest in Ireland. Additionally, species
648 identification errors and cases of synonymy lead to false negatives or cases of multiple
649 affiliations.

650

651 **Conclusion**

652 By comparing the accuracy and detection capacity of three metabarcoding strategies,
653 this study contributes to improving our toolkit for monitoring non-native insect
654 invasion. All three sequencing technologies performed equally well and showed similar
655 results for detecting and identifying exotic Cerambycid species collected in field traps.
656 However, the MinION stands out as a portable, easy-to-use, and cost-effective sequencer,
657 with the potential to become an essential tool for biodiversity monitoring projects. Using
658 the MinION reduces the time spent on laboratory handling compared to Illumina and

659 eliminates the need to outsource sample sequencing. This saves considerable time when
660 it comes to detecting invasive species. The MinION technology is accurate enough to
661 detect non-native species even when present at low abundances in field traps and allows
662 for accurate identifications as long as there is a sufficiently complete high-quality
663 reference database to avoid identification errors or false positives/negatives. It is also
664 crucial to pay close attention to issues of contamination and insect preservation during
665 and after individual capture in order to work with the least degraded DNA possible.

666

667 **References**

668 Abeynayake SW, Fiorito S, Dinsdale A, Whattam M, Crowe B, Sparks K, Campbell PR,
669 Gambley C (2021) A Rapid and Cost-Effective Identification of Invertebrate Pests at the
670 Borders Using MinION Sequencing of DNA Barcodes. *Genes* 12: 1138.
671 <https://doi.org/10.3390/genes12081138>

672

673 Alberdi A, Aizpurua O, Gilbert MTP, Bohmann K (2018) Scrutinizing key steps for reliable
674 metabarcoding of environmental samples. Mahon A (Ed.). *Methods in Ecology and*
675 *Evolution* 9: 134–147. <https://doi.org/10.1111/2041-210X.12849>

676

677 Andújar C, Arribas P, Gray C, Bruce C, Woodward G, Yu DW, Vogler AP (2018)
678 Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill.
679 *Molecular Ecology* 27: 146–166. <https://doi.org/10.1111/mec.14410>

680

681 Aukema JE, McCullough DG, Von Holle B, Liebhold AM, Britton K, Frankel SJ (2010)
682 Historical Accumulation of Nonindigenous Forest Pests in the Continental United States.
683 *BioScience* 60: 886–897. <https://doi.org/10.1525/bio.2010.60.11.5>

684

685 Ballare KM, Pope NS, Castilla AR, Cusser S, Metz RP, Jha S (2019) Utilizing field collected
686 insects for next generation sequencing: Effects of sampling, storage, and DNA extraction
687 methods. *Ecology and Evolution* 9: 13690–13705. <https://doi.org/10.1002/ece3.5756>

688

689 Batovska J, Piper AM, Valenzuela I, Cunningham JP, Blacket MJ (2021) Developing a non-
690 destructive metabarcoding protocol for detection of pest insects in bulk trap catches.
691 *Scientific Reports* 11: 7946. <https://doi.org/10.1038/s41598-021-85855-6>

692

693 Bensasson D, Zhang D-X, Hartl DL, Hewitt GM (2001) Mitochondrial pseudogenes:
694 evolution's misplaced witnesses. *Trends in Ecology & Evolution* 16: 314–321.
695 [https://doi.org/10.1016/s0169-5347\(01\)02151-6](https://doi.org/10.1016/s0169-5347(01)02151-6)

696

697 Blackburn TM, Pyšek P, Bacher S, Carlton JT, Duncan RP, Jarošík V, Wilson JRU,
698 Richardson DM (2011) A proposed unified framework for biological invasions. *Trends in*
699 *Ecology & Evolution* 26: 333–339. <https://doi.org/10.1016/j.tree.2011.03.023>

700

701 Bradshaw CJA, Leroy B, Bellard C, Roiz D, Albert C, Fournier A, Barbet-Massin M, Salles J-
702 M, Simard F, Courchamp F (2016) Massive yet grossly underestimated global costs of
703 invasive insects. *Nature Communications* 7: 12986.
704 <https://doi.org/10.1038/ncomms12986>

705

706 Braukmann TWA, Ivanova NV, Prosser SWJ, Elbrecht V, Steinke D, Ratnasingham S, De
707 Waard JR, Sones JE, Zakharov EV, Hebert PDN (2019) Metabarcoding a diverse arthropod
708 mock community. *Molecular Ecology Resources* 19: 711–727.
709 <https://doi.org/10.1111/1755-0998.13008>

710

711 Brown SDJ, Collins RA, Boyer S, Lefort M, Malumbres-Olarte J, Vink CJ, Cruickshank RH
712 (2012) S PIDER: An R package for the analysis of species identity and evolution, with
713 particular reference to DNA barcoding. *Molecular Ecology Resources* 12: 562–565.
714 <https://doi.org/10.1111/j.1755-0998.2011.03108.x>

715

716 Chain FJJ, Brown EA, MacIsaac HJ, Cristescu ME (2016) Metabarcoding reveals strong
717 spatial structure and temporal turnover of zooplankton communities among marine and
718 freshwater ports. Cowie R (Ed.). *Diversity and Distributions* 22: 493–504.
719 <https://doi.org/10.1111/ddi.12427>

720

721 Chown SL, Hodgins KA, Griffin PC, Oakeshott JG, Byrne M, Hoffmann AA (2015)
722 Biological invasions, climate change and genomics. *Evolutionary Applications* 8: 23–46.
723 <https://doi.org/10.1111/eva.12234>

724

725 Chua PYS, Bourlat SJ, Ferguson C, Korlevic P, Zhao L, Ekrem T, Meier R, Lawniczak MKN
726 (2023) Future of DNA-based insect monitoring. *Trends in Genetics* 39: 531–544.
727 <https://doi.org/10.1016/j.tig.2023.02.012>

728

729 Collins RA, Wangenstein OS, O’Gorman EJ, Mariani S, Sims DW, Genner MJ (2018)
730 Persistence of environmental DNA in marine systems. *Communications Biology* 1: 185.
731 <https://doi.org/10.1038/s42003-018-0192-6>

732

733 Dean MD, Ballard JWO (2001) Factors affecting mitochondrial DNA quality from museum
734 preserved *Drosophila simulans*. *Entomologia Experimentalis et Applicata* 98: 279–283.
735 <https://doi.org/10.1046/j.1570-7458.2001.00784.x>

736

737 deWaard JR, Levesque-Beaudin V, deWaard SL, Ivanova NV, McKeown JTA, Miskie R, Naik
738 S, Perez KHJ, Ratnasingham S, Sobel CN, Sones JE, Steinke C, Telfer AC, Young AD, Young
739 MR, Zakharov EV, Hebert PDN (2019) Expedited assessment of terrestrial arthropod
740 diversity by coupling Malaise traps with DNA barcoding 1. *Genome* 62: 85–95.
741 <https://doi.org/10.1139/gen-2018-0093>

742

743 Dopheide A, Tooman LK, Grosser S, Agabiti B, Rhode B, Xie D, Stevens MI, Nelson N,
744 Buckley TR, Drummond AJ, Newcomb RD (2019) Estimating the biodiversity of
745 terrestrial invertebrates on a forested island using DNA barcodes and metabarcoding
746 data. *Ecological Applications* 29: 1–14. Available from:
747 <https://www.jstor.org/stable/26669248> (June 9, 2023).

748

749 Egeter B, Peixoto S, Brito JC, Jarman S, Puppo P, Velo-Antón G (2018) Challenges for
750 assessing vertebrate diversity in turbid Saharan water-bodies using environmental DNA.
751 *Genome* 61: 807–814. <https://doi.org/10.1139/gen-2018-0071>

752

753 Egeter B, Veríssimo J, Lopes-Lima M, Chaves C, Pinto J, Riccardi N, Beja P, Fonseca NA
754 (2022) Speeding up the detection of invasive bivalve species using environmental DNA:
755 A Nanopore and Illumina sequencing comparison. *Molecular Ecology Resources* 22:
756 2232–2247. <https://doi.org/10.1111/1755-0998.13610>

757

758 Elbrecht V, Leese F (2017) Validation and Development of COI Metabarcoding Primers
759 for Freshwater Macroinvertebrate Bioassessment. *Frontiers in Environmental Science* 5.
760 Available from: <https://www.frontiersin.org/articles/10.3389/fenvs.2017.00011> (July 2,
761 2023).

762

763 Elbrecht V, Vamos EE, Meissner K, Aroviita J, Leese F (2017) Assessing strengths and
764 weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine
765 stream monitoring. Yu D (Ed.). *Methods in Ecology and Evolution* 8: 1265–1275.
766 <https://doi.org/10.1111/2041-210X.12789>

767

768 Elbrecht V, Braukmann TWA, Ivanova NV, Prosser SWJ, Hajibabaei M, Wright M, Zakharov
769 EV, Hebert PDN, Steinke D (2019) Validation of COI metabarcoding primers for
770 terrestrial arthropods. *PeerJ* 7: e7745. <https://doi.org/10.7717/peerj.7745>

771

772 Escudié F, Auer L, Bernard M, Mariadassou M, Cauquil L, Vidal K, Maman S, Hernandez-
773 Raquet G, Combes S, Pascal G (2018) FROGS: Find, Rapidly, OTUs with Galaxy Solution.
774 Berger B (Ed.). *Bioinformatics* 34: 1287–1294.
775 <https://doi.org/10.1093/bioinformatics/btx791>

776

777 Ferro ML, Park J-S (2013) Effect of Propylene Glycol Concentration on Mid-Term DNA
778 Preservation of Coleoptera. In: The Coleopterists Bulletin. , 581–586.
779 <https://doi.org/10.1649/0010-065X-67.4.581>

780

781 Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-
782 generation sequencing data. Bioinformatics (Oxford, England) 28: 3150–3152.
783 <https://doi.org/10.1093/bioinformatics/bts565>

784

785 Gibson J, Shokralla S, Porter TM, King I, van Konynenburg S, Janzen DH, Hallwachs W,
786 Hajibabaei M (2014) Simultaneous assessment of the macrobiome and microbiome in a
787 bulk sample of tropical arthropods through DNA metasytematics. Proceedings of the
788 National Academy of Sciences of the United States of America 111: 8007–8012.
789 <https://doi.org/10.1073/pnas.1406468111>

790

791 Giovani B, Blümel S, Lopian R, Teulon D, Bloem S, Galeano Martínez C, Beltrán Montoya C,
792 Urias Morales CR, Dharmapuri S, Timote V, Horn N, Chouibani M, Mezui M’Ella JG,
793 Herrera V, Castinel A, Goletsos C, Moeller C, Naumann I, Stancanelli G, Bronzwaer S,
794 Tramontini S, MacDonald P, Matheson L, Anthoine G, De Jonghe K, Schenk M, Steinmüller
795 S, Rodriguez E, Cruz ML, Luck J, Fraser G, Brunel S, Montuori M, Fedchock C, Steel E,
796 Pennington HG, Day R, Rossi JP, Xia J (2020) Science diplomacy for plant health. Nature
797 Plants 6: 902–905. <https://doi.org/10.1038/s41477-020-0744-x>

798

799 Haran J, Koutroumpa F, Magnoux E, Roques A, Roux G (2015) Ghost mtDNA haplotypes
800 generated by fortuitous NUMTs can deeply disturb infra-specific genetic diversity and

801 phylogeographic pattern. *Journal of Zoological Systematics and Evolutionary Research*
802 53: 109–115. <https://doi.org/10.1111/jzs.12095>

803

804 Harrison JB, Sunday JM, Rogers SM (2019) Predicting the fate of eDNA in the
805 environment and implications for studying biodiversity. *Proceedings of the Royal Society*
806 B: Biological Sciences 286: 20191409. <https://doi.org/10.1098/rspb.2019.1409>

807

808 Hebert PDN, Ratnasingham S, deWaard JR (2003) Barcoding animal life: cytochrome c
809 oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal*
810 *Society B: Biological Sciences* 270: S96–S99. <https://doi.org/10.1098/rsbl.2003.0025>

811

812 Hebert PDN, deWaard JR, Zakharov EV, Prosser SWJ, Sones JE, McKeown JTA, Mantle B,
813 La Salle J (2013) A DNA ‘Barcode Blitz’: Rapid Digitization and Sequencing of a Natural
814 History Collection. Kolokotronis S-O (Ed.). *PLoS ONE* 8: e68535.
815 <https://doi.org/10.1371/journal.pone.0068535>

816

817 Henrie JR, Thomson BM, Yungfleisch AA, Kerr M, Leavitt SD (2022) Characterizing
818 Crustose Lichen Communities—DNA Metabarcoding Reveals More than Meets the Eye.
819 *Diversity* 14: 766. <https://doi.org/10.3390/d14090766>

820

821 Ho JKI, Puniamoorthy J, Srivathsan A, Meier R (2020) MinION sequencing of seafood in
822 Singapore reveals creatively labelled flatfishes, confused roe, pig DNA in squid balls, and
823 phantom crustaceans. *Food Control* 112: 107144.
824 <https://doi.org/10.1016/j.foodcont.2020.107144>

825

826 Hodgetts J, Ostojá-Starzewski JC, Prior T, Lawson R, Hall J, Boonham N (2016) DNA
827 barcoding for biosecurity: case studies from the UK plant protection program. Wilson J-J
828 (Ed.). *Genome* 59: 1033–1048. <https://doi.org/10.1139/gen-2016-0010>

829

830 Javal M, Terblanche J, Conlong D, Delahaye N, Grobbelaar E, Benoit L, Lopez-Vaamonde C,
831 Haran J (2021) DNA barcoding for bio-surveillance of emerging pests and species
832 identification in Afrotropical Prioninae (Coleoptera, Cerambycidae). *Biodiversity Data*
833 *Journal* 9: e64499. <https://doi.org/10.3897/BDJ.9.e64499>

834

835 Kelnarova I, Jendek E, Grebennikov VV, Bocak L (2019) First molecular phylogeny of
836 *Agrilus* (Coleoptera: Buprestidae), the largest genus on Earth, with DNA barcode
837 database for forestry pest diagnostics. *Bulletin of Entomological Research* 109: 200–211.
838 <https://doi.org/10.1017/S0007485318000330>

839

840 Kenis M, Auger-Rozenberg M-A, Roques A, Timms L, Péré C, Cock MJW, Settele J, Augustin
841 S, Lopez-Vaamonde C (2009) Ecological effects of invasive alien insects. *Biological*
842 *Invasions* 11: 21–45. <https://doi.org/10.1007/s10530-008-9318-y>

843

844 Kitson JJN, Hahn C, Sands RJ, Straw NA, Evans DM, Lunt DH (2019) Detecting host-
845 parasitoid interactions in an invasive Lepidopteran using nested tagging DNA
846 metabarcoding. *Molecular Ecology* 28: 471–483. <https://doi.org/10.1111/mec.14518>

847

848 Koutroumpa FA, Lieutier F, Roux-Morabito G (2009) Incorporation of mitochondrial
849 fragments in the nuclear genome (*Numts*) of the longhorned beetle *Monochamus*
850 *galloprovincialis* (Coleoptera, Cerambycidae). Journal of Zoological Systematics and
851 Evolutionary Research 47: 141–148. [https://doi.org/10.1111/j.1439-](https://doi.org/10.1111/j.1439-0469.2008.00492.x)
852 [0469.2008.00492.x](https://doi.org/10.1111/j.1439-0469.2008.00492.x)

853

854 Krehenwinkel H, Pomerantz A, Prost S (2019) Genetic Biomonitoring and Biodiversity
855 Assessment Using Portable Sequencing Technologies: Current Uses and Future
856 Directions. Genes 10: 858. <https://doi.org/10.3390/genes10110858>

857

858 Lindahl, T. "Instability and Decay of the Primary Structure of DNA." Nature, vol. 362, no.
859 6422, Apr. 1993, pp. 709–15, <https://doi.org/10.1038/362709a0>.

860

861 Liu M, Clarke LJ, Baker SC, Jordan GJ, Burrige CP (2020) A practical guide to DNA
862 metabarcoding for entomological ecologists. Ecological Entomology 45: 373–385.
863 <https://doi.org/10.1111/een.12831>

864

865 Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M (2014) Swarm: robust and fast
866 clustering method for amplicon-based studies. PeerJ 2: e593.
867 <https://doi.org/10.7717/peerj.593>

868

869 Martin M (2011) Cutadapt removes adapter sequences from high-throughput
870 sequencing reads. EMBnet.journal 17: 10. <https://doi.org/10.14806/ej.17.1.200>

871

872 Mas H, Santoiemma G, Lencina JL, Gallego D, Pérez-Laorga E, Ruzzier E, Rassati D (2023)
873 Investigating beetle communities in and around entry points can improve surveillance at
874 national and international scale. *NeoBiota* 85: 145–165.
875 <https://doi.org/10.3897/neobiota.85.103904>

876

877 Mazza G, Tricarico E, Genovesi P, Gherardi F (2014) Biological invaders are threats to
878 human health: an overview. *Ethology Ecology & Evolution* 26: 112–129.
879 <https://doi.org/10.1080/03949370.2013.863225>

880

881 Meusnier, Isabelle, et al. "A Universal DNA Mini-Barcode for Biodiversity Analysis." *BMC*
882 *Genomics*, vol. 9, no. 1, May 2008, p. 214, <https://doi.org/10.1186/1471-2164-9-214>.

883 Piper AM, Batovska J, Cogan NOI, Weiss J, Cunningham JP, Rodoni BC, Blacket MJ (2019)
884 Prospects and challenges of implementing DNA metabarcoding for high-throughput
885 insect surveillance. *GigaScience* 8: giz092. <https://doi.org/10.1093/gigascience/giz092>

886

887 Preston, M., Fritzsche, M. & Woodcock, P. 2022. Understanding and mitigating errors and
888 biases in metabarcoding: an introduction for non-specialists. JNCC Report No. 699, JNCC,
889 Peterborough, ISSN 0963-8091.

890

891 Ratnasingham S, Hebert PDN (2007) BARCODING: bold: The Barcode of Life Data System
892 (<http://www.barcodinglife.org>): BARCODING. *Molecular Ecology Notes* 7: 355–364.
893 <https://doi.org/10.1111/j.1471-8286.2007.01678.x>

894

895 Richardson DM, Pysek P, Rejmanek M, Barbour MG, Panetta FD, West CJ (2000)
896 Naturalization and invasion of alien plants: concepts and definitions. *Diversity*
897 <https://doi.org/10.1046/j.1472-4642.2000.00083.x>
898 <https://doi.org/10.1046/j.1472-4642.2000.00083.x>

899
900 Roques A, Auger-Rozenberg M-A, Blackburn TM, Garnas J, Pyšek P, Rabitsch W,
901 Richardson DM, Wingfield MJ, Liebhold AM, Duncan RP (2016) Temporal and
902 interspecific variation in rates of spread for insect species invading Europe during the
903 last 200 years. *Biological Invasions* 18: 907–920. [https://doi.org/10.1007/s10530-016-](https://doi.org/10.1007/s10530-016-1080-y)
904 [1080-y](https://doi.org/10.1007/s10530-016-1080-y)

905
906 Roques A, Ren L, Rassati D, Shi J, Akulov E, Audsley N, Auger-Rozenberg M-A, Avtzis D,
907 Battisti A, Bellanger R, Bernard A, Bernadinelli I, Branco M, Cavaletto G, Cocquempot C,
908 Contarini M, Courtial B, Courtin C, Denux O, Dvořák M, Fan J, Feddern N, Francese J,
909 Franzen EKL, Garcia A, Georgiev G, Georgieva M, Giarruzzo F, Gossner M, Gross L,
910 Guarneri D, Hoch G, Hölling D, Jonsell M, Kirichenko N, Loomans A, Luo Y, McCullough D,
911 Maddox C, Magnoux E, Marchioro M, Martinek P, Mas H, Mériguet B, Pan Y, Phélut R,
912 Pineau P, Ray AM, Roques O, Ruiz M-C, Monteys VS i, Speranza S, Sun J, Sweeney JD,
913 Touroult J, Valladares L, Veillat L, Yuan Y, Zalucki MP, Zou Y, Žunič-Kosi A, Hanks LM,
914 Millar JG (2023) Worldwide tests of generic attractants, a promising tool for early
915 detection of non-native cerambycid species. *NeoBiota* 84: 169–209.
916 <https://doi.org/10.3897/neobiota.84.91096>

917

918 Rossa R, Goczał J (2021) Global diversity and distribution of longhorn beetles
919 (Coleoptera: Cerambycidae). *The European Zoological Journal* 88: 289–302.
920 <https://doi.org/10.1080/24750263.2021.1883129>

921
922 Sahlin K, Medvedev P (2020) De Novo Clustering of Long-Read Transcriptome Data
923 Using a Greedy, Quality Value-Based Algorithm. *Journal of Computational Biology: A*
924 *Journal of Computational Molecular Cell Biology* 27: 472–484.
925 <https://doi.org/10.1089/cmb.2019.0299>

926
927 Seebens H, Bacher S, Blackburn TM, Capinha C, Dawson W, Dullinger S, Genovesi P,
928 Hulme PE, Kleunen M, Kühn I, Jeschke JM, Lenzner B, Liebhold AM, Pattison Z, Pergl J,
929 Pyšek P, Winter M, Essl F (2021) Projecting the continental accumulation of alien species
930 through to 2050. *Global Change Biology* 27: 970–982.
931 <https://doi.org/10.1111/gcb.15333>

932
933 Seebens H, Blackburn TM, Dyer EE, Genovesi P, Hulme PE, Jeschke JM, Pagad S, Pyšek P,
934 van Kleunen M, Winter M, Ansong M, Arianoutsou M, Bacher S, Blasius B, Brockerhoff EG,
935 Brundu G, Capinha C, Causton CE, Celesti-Grapow L, Dawson W, Dullinger S, Economo EP,
936 Fuentes N, Guénard B, Jäger H, Kartesz J, Kenis M, Kühn I, Lenzner B, Liebhold AM,
937 Mosena A, Moser D, Nentwig W, Nishino M, Pearman D, Pergl J, Rabitsch W, Rojas-
938 Sandoval J, Roques A, Rorke S, Rossinelli S, Roy HE, Scalera R, Schindler S, Štajerová K,
939 Tokarska-Guzik B, Walker K, Ward DF, Yamanaka T, Essl F (2018) Global rise in emerging
940 alien species results from increased accessibility of new source pools. *Proceedings of the*
941 *National Academy of Sciences* 115. <https://doi.org/10.1073/pnas.1719429115>

942

943 Silvestre-Ryan J, Holmes I (2021) Pair consensus decoding improves accuracy of neural
944 network basecallers for nanopore sequencing. *Genome Biology* 22: 38.
945 <https://doi.org/10.1186/s13059-020-02255-1>

946

947 Srivathsan, Amrita, et al. "ONTbarcoder 2.0: Rapid Species Discovery and Identification
948 with Real-time Barcoding Facilitated by Oxford Nanopore R10 .4." *Cladistics*, vol. 40, no.
949 2, 2024, pp. 192–203, <https://doi.org/10.1111/cla.12566>.

950

951 Srivathsan A, Lee L, Katoh K, Hartop E, Kutty SN, Wong J, Yeo D, Meier R (2021)
952 ONTbarcoder and MinION barcodes aid biodiversity discovery and identification by
953 everyone, for everyone. *BMC Biology* 19: 217. [https://doi.org/10.1186/s12915-021-](https://doi.org/10.1186/s12915-021-01141-x)
954 [01141-x](https://doi.org/10.1186/s12915-021-01141-x)

955

956 Strickler KM, Fremier AK, Goldberg CS (2015) Quantifying effects of UV-B, temperature,
957 and pH on eDNA degradation in aquatic microcosms. *Biological Conservation* 183: 85–
958 92. <https://doi.org/10.1016/j.biocon.2014.11.038>

959

960 Vamos E, Elbrecht V, Leese F (2017) Short COI markers for freshwater macroinvertebrate
961 metabarcoding. *Metabarcoding and Metagenomics* 1: e14625.
962 <https://doi.org/10.3897/mbmg.1.14625>

963

964 Vaser R, Sović I, Nagarajan N, Šikić M (2017) Fast and accurate de novo genome assembly
965 from long uncorrected reads. *Genome Research* 27: 737–746.
966 <https://doi.org/10.1101/gr.214270.116>

967

968 Wang Q (2017) *Cerambycidae of the World: Biology and Pest Management*. CRC Press,
969 643 pp.

970

971 Wu Y, Trepanowski NF, Molongoski JJ, Reagel PF, Lingafelter SW, Nadel H, Myers SW, Ray
972 AM (2017) Identification of wood-boring beetles (Cerambycidae and Buprestidae)
973 intercepted in trade-associated solid wood packaging material using DNA barcoding and
974 morphology. *Scientific Reports* 7: 40316. <https://doi.org/10.1038/srep40316>

975

976 Zenker MM, Rougerie R, Teston JA, Laguerre M, Pie MR, Freitas AVL (2016) Fast Census
977 of Moth Diversity in the Neotropics: A Comparison of Field-Assigned Morphospecies and
978 DNA Barcoding in Tiger Moths. Cimmaruta R (Ed.). *PLOS ONE* 11: e0148423.
979 <https://doi.org/10.1371/journal.pone.0148423>

980

981 **Acknowledgements**

982 We would like to thank all colleagues who participated in the taxa sampling (see Roques
983 et al. 2023). We are thankful to Lucas Sire for insightful discussion on primer choice. The
984 authors thank the GeT-Biopuces platform of INSA Toulouse for the IonTorrent
985 sequencing study. This work was supported by the PORTRAP project “Test de l’efficacité
986 de pièges génériques multicomposés pour la détection précoce d’insectes exotiques
987 xylophages dans les sites potentiels d’entrée sur le territoire national” and HOMED

988 project (HOListic Management of Emerging Forest Pests and Diseases) which received
989 funding from the European Union's Horizon 2020 research and innovation program
990 under grant agreement No. 771271 (<https://homed-projecteu/>). We are grateful to the
991 genotoul bioinformatics platform Toulouse Midi-Pyrenees for providing help and
992 computing storage resources. Loïs Veillat was supported by a PhD studentship from
993 HOMED project and doctoral school SSBCV at the university of Orléans.

994

995 **Data accessibility**

996 Barcode data for the 33 species used in the mock community experiment are available
997 from BOLD in the dataset DS-MINION (dx.doi.org/10.5883/DS-MINION). Raw sequence
998 data for this project and analytical script and files are available on figshare
999 ([https://figshare.com/projects/DNA_metabarcoding_an_efficient_way_to_detect_non-
1000 native_cerambycid_beetles_in_trapping_collections_/171432](https://figshare.com/projects/DNA_metabarcoding_an_efficient_way_to_detect_non-native_cerambycid_beetles_in_trapping_collections_/171432)).

1001

1002 **Author contributions**

1003 Loïs Veillat, Géraldine Roux, Carlos Lopez-Vaamonde and Stéphane Boyer conceived the
1004 study. Alain Roques collected field samples. Stéphane Boyer, Marina Querejeta,
1005 Emmanuelle Magnoux and Loïs Veillat conducted the laboratory sample processing. Loïs
1006 Veillat analysed the data and wrote the first draft. All authors contributed to the
1007 preparation of the manuscript. Both senior authors, Géraldine Roux and Carlos Lopez-
1008 Vaamonde, contributed equally to this study.