Climate-induced forest dieback drives compositional change in insect communities that is concentrated amongst rare species

Lucas Sire[1, *], Paul Schmidt Yáñez[2], Cai Wang[3,4], Annie Bézier[1], Béatrice Courtial[5], Jérémy Cours[6], Diego Fontaneto[7], Laurent Larrieu[8,9], Christophe Bouget[6], Simon Thorn[10], Jörg Müller[11,12], Douglas W. Yu[3,13], Michael T. Monaghan[2,14], Elisabeth A. Herniou[1] & Carlos Lopez-Vaamonde[1,5].

1: Institut de Recherche sur la Biologie de l'Insecte (IRBI), UMR 7261, CNRS-Université de Tours, Tours, France

2: Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), Müggelseedamm 301, 12587 Berlin, Germany

3: State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China 650223

4: Kunming College of Life Sciences, University of Chinese Academy of Sciences, Kunming, China

5: INRAE, Zoologie Forestière, F-45075 Orléans, France

6: INRAE 'Forest Ecosystems' Research Unit – Biodiversity team Domaine des Barres F-45290 Nogent-sur-Vernisson France

7: Water Research Institute, National Research Council of Italy, CNR-IRSA, Largo Tonolli 50, 28922 Verbania Pallanza, Italy

8: Université de Toulouse, INRAE, UMR DYNAFOR, Castanet-Tolosan, France

9: CRPF Occitanie, Tarbes, France

10: Field Station Fabrikschleichach, Biocenter, University of Würzburg, Glashüttenstraße. 5, 96181 Rauhenebrach, Germany

11: Field Station Fabrikschleichach, Department of Animal Ecology and Tropical Biology, Biocenter, University of Würzburg, Glashüttenstraße 5, 96181 Rauhenebrach, Germany

12: Bavarian Forest National Park, Freyunger Str. 2, 94481 Grafenau, Germany

13: School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk, UK NR47TJ

14: Institut für Biologie, Freie Universität Berlin, Königin-Luise-Straße. 1-3, 12489 Berlin, Germany

* corresponding author: lucas.sire@univ-tours.fr (*L. Sire*)

Abstract

Marked decline in insect species richness, abundance and biomass have recently been quantified in Europe. We metabarcoded 224 Malaise-trap samples to investigate whether drought-induced forest dieback and subsequent salvage logging have an impact on flying insects (*ca*. 3000 insect species) in silver fir Pyrenean forests. We found no evidence that climate-induced forest dieback impacted species richness of flying insects but revealed compositional turnover patterns consistent with those seen during natural forest succession, given that the key covariates explaining compositional variation were canopy openness versus microhabitat diversity and deadwood amount at local and landscape scales, respectively. Importantly, most change was driven by rare species. In contrast, observed levels of salvage logging did not explain change in species richness or composition. Hence, although forest dieback appears to cause changes in species assemblages mimicking natural forest succession, it also increases the risk of catastrophic loss of rare species through homogenization of environmental conditions.

Keywords

Abies alba, biomonitoring, DNA Metabarcoding, climate change, spatial and temporal species turnover, Malaise traps, salvage logging, zeta diversity, Pyrenees

Introduction

Insects are vital components of biodiversity, providing important ecosystem services such as pollination and pests regulation, while also performing disservices as disease vectors and plant pests[1]. Global changes, including those of climate, land use and land cover, can lead to degradation and habitat loss, chemical and light pollution or invasive species. These changes have caused major drops in biomass, abundance and species richness of insects[2], and this accelerating decline has become a major cause for concern. Yet, understanding the relative impacts of these drivers underlying insect decline and its propensity is complex[3].

Forests are thought to act as buffer zones from rapid anthropogenic changes in adjacent areas, and contribute to biodiversity conservation by providing refuge for rich insect communities[4]. However, forests are increasingly suffering from climate-induced tree dieback, regardless of protection measures[5], which could have long term consequences for forest biodiversity. Tree diebacks largely result from more frequent, intense and longer droughts, especially in Europe[6]. These can induce tree mortality especially if they promote insect pests and pathogen outbreaks. Some insect species benefit from warmer temperatures (*i.e.* mild winters) allowing distribution range expansion and population outbreaks, which can cause massive forest diebacks[7]. Bioclimatic models predict that the extent, severity and duration of droughts will increase as a result of climate change, notably in temperate and Mediterranean areas[8].

Tree diebacks can cause major structural changes such as an increase in canopy openness, reduced foliage density, which in turn increases light availability, potentially changing the community structure of understory plants and their associated herbivorous insects[9]. While insect decline is independent from forest protection status, it has been presented as lower in terms of species richness for plots undergoing dieback[2]. Indeed, tree dieback can also increase the availability of nutrients stored in vital trees (*i.e.* deadwood), sunlight and microhabitats, increasing species richness for multiple insect groups including bees, wasps, hoverflies, saproxylic beetles, as well as multiple red-listed insect taxa following insect pest outbreaks[10], or canopy-dwelling beetles in declining oak forests[9]. However, no significant change in species richness of ground-dwelling carabids and spiders has been associated with climate-induced dieback in a beech-dominated forest[11]. Meanwhile, forest management practices, like salvage logging (*i.e.* cutting down dying trees to salvage their timber value), can also have contrasting effects on biodiversity[12]. For instance, species richness of saproxylic insects increased with bark beetle outbreaks[10], but decreased with deadwood harvesting[13]. Even

3

88   though specific to each of these well-studied taxa previously stated, species richness in general

89   appears to be boosted by tree diebacks, while salvage logging seems to have no effect overall[14].

90   The identification of samples down to species level is important in ecological studies

91   because adaptive response to disturbances can be highly species-specific[15]. Moreover, these

92   studies on insect biodiversity response to forest disturbances are mainly based on species

93   richness, which has been used as a surrogate for ecosystem functionality[16]. However, richness

94   alone is often a poor indicator of biodiversity change compared to ecological guilds changes to

95   the community[17]. Quantifying these changes requires high taxonomic resolution that is often

96   not available. The morphological identification that richness measures are typically derived

97   from are often limited to well-known insect bioindicators that are overrepresented in the

98   literature, hence biasing species richness *per se* and impeding further holistic community-based

99   and ecological studies. Unfortunately, accurate species identification of hyper-diverse insect

100   groups such as dipterans and hymenopterans is difficult because of the taxonomic

101   impediment[18], often hampering—or limiting to few remarkable groups only—the study of these

102   species-rich taxa despite their ecological importance[19]. The use of DNA barcoding as a tool for

103   species identification[20] now allows to bypass this taxonomic impediment[21]. As exemplified by

104   Wang *et al.*[22] targeting poorly described Chinese entomofauna, the use of DNA metabarcoding

105   allows to tackle more comprehensive insect biodiversity studies and facilitates the documenting

106   of all trophic guilds and their responses to forest disturbances. The authors suggested that forest

107   diebacks induced by bark beetle outbreaks could drive a transition from homogenous

108   plantations to more biodiversity-friendly heterogenous forests[22].

109   The response to forest dieback and subsequent salvage logging varies among taxa and

110   functional groups. Floricolous species are expected to increase in diversity and abundance from

111   tree dieback due to greater canopy openness[15]. Similarly, parasitoids may benefit from

112   temperature rises and associated tree diebacks as the more complex a forest is the more diverse

113   the parasitoid communities are[23], but may also suffer from poorer host quality, especially of

114   sap-feeders directly affected by tree health under drought conditions[24]. Forests undergoing

115   diebacks are thus expected to induce guild-specific responses, which may have contrasting

116   effects on species richness of each insect groups depending on their respective ecology. A

117   comprehensive approach on sampling and identification of taxa is therefore needed to improve

118   our understanding of the effects of forest disturbances on insect biodiversity.

119   Here, we study insect diversity in montane Pyrenean forests dominated by silver fir

120   (*Abies alba* Mill), a conifer species sensitive to drought[25] that suffered severe climate-induced

121   diebacks after a heatwave in 2003[26]. Currently, forest management practices often implement

salvage logging at the first signs of dieback[26]. Despite the high conservation value of Pyrenean silver fir forests, the effects of dieback and salvage logging on associated insect fauna remains unknown. To address this gap, we studied the communities of aerial insects over 56 silver fir forest plots (Supplementary Fig. 1) that varied in dieback level and management practices. From late spring to early autumn 2017, insects were mass-trapped monthly using Malaise traps, resulting in 224 samples that we analyzed using DNA metabarcoding of a fragment of the mitochondrial COI gene region. Based on BOLD DNA reference libraries[27], taxonomy of the recovered Molecular Operational Taxonomic Units (MOTUs) was assessed and broad ecological functions (*i.e.* floricolous and parasitoid insects) were attributed using family-level information from published literature whenever possible. We used generalized linear models to assess whether dieback levels and salvage logging influence the structure and diversity of forest insect communities and functional guilds. We hypothesized that levels of forest dieback and salvage logging change species composition of local insect communities, which is seen as turnover in community structure across sites with different disturbances. We also hypothesized that functional guilds respond differently to forest dieback intensity; in particular the diversity of parasitoids and floricolous insects in areas of high forest dieback. As we found no change in species richness but variations in community compositions across dieback levels and management practices, we performed zeta order analyses[28,29] to further investigate the dataset structure and identify the nature of these community changes, as well as the species compositional and functional turnover. We also used zeta analyses in multi-site generalized dissimilarity modelling (zeta.msgdm)[30] to study the impact of environmental features (*i.e.* geographic distance to the nearest plot, altitude, canopy openness, total amount of deadwood, basal area, tree diversity, density of very large trees and Tree-related Microhabitats (TreM) diversity) in driving these community structural changes and further discuss their associated consequences for forest management and conservation. Finally, we used the fourth-corner model[31] to highlight winning and losing insects to forest disturbances as well as idiosyncratic responses of insect orders and functional groups.

Results

Diverse insect community with high temporal and spatial turnover

Metabarcoding recovered thousands of locally rare species and unnamed "dark taxa" (*i.e.* specie-level 'invertebrate' taxa in DNA reference libraries without a species name assigned and/or species-rich taxa, often small in body size and under-described)[32]. From the 224 Malaise trap samples, we recovered 2972 MOTUs, with iNEXT extrapolation estimating a total richness of ~4000 MOTUs and species accumulation curves reaching a plateau (Fig. 1A, Supplementary Fig. 2). Our 4-month sampling effort (from late spring to early autumn) thus captured *ca.* 75% of the total Malaise-trappable insect diversity in the sampled areas allowing us to reach ~90% sample coverage (Fig. 1A–C). Applying a threshold of 97% sequence similarity for species-level discrimination and keeping unambiguous taxonomic matches for higher taxonomic ranks, 100%, 84.8%, 75.4% and 52.4% (1558) of the total 2972 insect MOTUs could be assigned to order, family, genus and species names, respectively (Supplementary Table I). Two orders, Diptera and Hymenoptera, together represented 73% of all MOTUs, as expected with Malaise-trapped samples (Fig. 2A), with 641 (49.7%) and 371 MOTUs (41.4%) identified to species, respectively (Supplementary Fig. 4). As with the whole dataset (Fig. 1A), the accumulation curves of the five most represented insect orders approached asymptotes (Fig. 2B).

We observed a high rate of temporal species turnover across months, as evidenced by only 360 MOTUs (12% of the total MOTUs diversity) being detected in all fourth months (Fig. 1D). Thus, even though our sampling was efficient at recovering high biodiversity, sampling throughout the complete growing season is likely to increase the total diversity occurring in the region (Fig. 1D). Geographic turnover was also observed, with only 45% of the total insect species sampled occurring in both of the sampled districts Aure Valley (Central French Pyrenees) and Sault Plateau (Eastern French Pyrenees) (Supplementary Fig. 1), the remaining 31% and 24% being specific to each, respectively. Additionally, these district-specific taxa were the most rarely caught throughout all the different plots of their respective districts (Supplementary Fig. 3A).

Forest disturbances influences the composition but not the species richness of insect communities

We grouped sites into three dieback categories (low, medium, high) based on the proportion of drought-affected and dying trees for each plot. Insect species richness was

6

compared across the three dieback categories but also across three defined stand types including management practice (healthy, stands expressing dieback but not salvaged—hereafter 'disturbed'—, stands expressing dieback and being salvage logged—hereafter shortened as salvaged). No response could be detected in terms of changes in species richness from all insect or functional groups tested, neither for the three levels of forest dieback nor the stand types (Table I). However, a significant difference in species richness was found between the two sampled districts for Coleoptera only, with more species of beetles detected in Sault plateau (Table I).

As the dataset was geographically structured (*i.e.* Aure valley and Sault plateau districts), we analysed the dataset using the nearest-neighbour (NN) sample-selection scheme (Fig. 3). The NN model fitted better to a power-law function than to an exponential one (Fig. 3C, D; $AIC_{(NN, Exp)}$ = 0.89, $AIC_{(NN, PL)}$ = -122.78), which was consistent with community assembly being driven by niche differentiation over stochastic assembly. We also ran a zeta diversity analysis with the non-geographically structured all-combinations (ALL) sample-selection scheme (Supplementary Fig. 5), and the results were similar but showed a much weaker fit to the power-law function (Supplementary Fig. 5C, D; $AIC_{(ALL, Exp)}$ = 0.63, $AIC_{(ALL, PL)}$ = -20.97). The rapid decline in zeta diversity between zeta orders 2 and 10 indicated that compositional turnover was mainly driven by rare species; few species were shared in 10 or more sites (Fig. 3A; Supplementary Fig. 3B). However, re-visualising the decline curve in Fig. 3A as a zeta retention-rate curve (Fig. 3B) showed that the few species that did occur in ≥10 sites (n = 42 species) were highly prevalent but only two species were being shared by 55 out of 56 total sites. Interestingly, this decline curve also demonstrated the drop of species retention rate at 28 sites—equivalent to the number of sampled plots in each district—, accounting for the strong geographic effect that induced very few shared species between Aure valley and Sault plateau. Re-analysis using the ALL sample-selection scheme produced similar results in shared species but smoothed geographic effects (Supplementary Fig. 5).

The impacts of drought-induced forest dieback intensities and stand types on community compositional changes were evaluated for the total sampling (Fig. 4A). Overall, forest dieback was found to induce significant changes in insect community assemblages for all tested groups but Coleoptera (Supplementary Table II). However, no significant variation in community composition was found across stand types, hence no effect of salvage logging could be detected (Supplementary Table II). Regardless of dieback intensity and stand type, each insect community assemblage tested significantly differed between districts (Fig. 4B, Supplementary Table II). Each dieback category hosted particular sets of species, yet insect

7

216   communities of low dieback level plots were more similar to each other and more distinct from
217   those of medium and high dieback level plots (Fig. 4A; Supplementary Fig. 3B). Furthermore,
218   taxa specific to a particular dieback level were mostly rare taxa (*i.e.* taxa with low prevalence)
219   (Supplementary Fig. 3B). Finally, community composition variations found across district were
220   reflected by Sault plateau plots sharing more similar species than those of Aure valley plots
221   (Fig. 4B).

222

223   *Winners and losers of forest disturbances*

224   Zeta decline analyses allowed us to assess community assemblages and species
225   retention rates across plots for the five main insect orders represented and for functional groups
226   (floricolous / non-floricolous and parasitoids / non-parasitoid species) within each dieback
227   category and stand type. We detected that species were retained differently within both dieback
228   intensity gradient and stand types, regardless of the model scheme used (Fig. 5–6;
229   Supplementary Fig. 6–7). Both Lepidoptera and Hemiptera showed a quick drop and a lack of
230   structure in zeta ratio along the dieback gradient (Fig. 5A–C) and between the different stand
231   types (Fig. 6A–C), indicating no species shared across zeta order ranges of both environmental
232   gradients. Similar zeta declines were observed at higher zeta orders for Coleoptera at low
233   dieback level and in healthy stands (Fig. 5A; Fig. 6A), as well as for Diptera and Hymenoptera
234   at high dieback levels and in disturbed stands (Fig. 5C; Fig. 6B). These results highlighted that
235   rare species shaped Coleoptera communities mostly at low dieback level, whereas common
236   Coleoptera species could be found across plots of higher dieback gradients. Conversely, both
237   Diptera and Hymenoptera species assemblages were less diverse at low dieback level and
238   healthy stands while high dieback level and disturbed stands favoured complete species
239   turnover with no common species retained (Fig. 5A–C, Fig. 6A–B). Regarding functional
240   assemblages, all had common species likely to be found across low and medium dieback
241   gradient, as well as healthy and salvaged forest stands (Fig. 5D–E; Fig. 6D–F). Nevertheless,
242   while a similar pattern was observed for non floricolous and non-parasitoid species assemblages
243   at high dieback level or within disturbed stands, zeta ratio of decline for both parasitoids and
244   floricolous species cohorts fully dropped, indicating species compositional turnover within the
245   two functional groups and no core species shared across 18 to 23 zeta order range (Fig. 5F; Fig.
246   6E). This lack of structure was likely driven by effects of high dieback level and disturbed
247   stands observed on both Diptera and Hymenoptera, which included many taxa of pollinators
248   and/or parasitoids (Fig. 5C, Fig. 6B). Overall, while few drops in species retention rates due to
249   geographic effect were noticeable throughout the different environmental conditions (Fig. 5–

8

250   6), main effects of both dieback level and stand types on species retention rates remained visible

251   in ALL model scheme (Supplementary Fig. 6–7).

252   Winning and losing insect orders in terms of prevalence over dieback gradient and stand

253   types were assessed using a "fourth corner" modelling. Thus, we highlighted a higher

254   prevalence in low dieback level stands and conversely, a detrimental effect of salvage logging

255   over Coleoptera (Fig. 7). Furthermore, the lack of drop in species retention rate for both

256   Hymenoptera and Diptera at low dieback level or in healthy stands (Fig. 5A, Fig. 6A) could be

257   explained by a lower prevalence of these two orders in these environments (Fig. 7).

258   Interestingly, Hymenoptera but not Diptera were winning from a particular level of dieback,

259   especially with a higher prevalence at medium dieback level (Fig. 7A), while both Lepidoptera

260   and Hemiptera diversity were favoured by stand disturbances in general (Fig. 7B).

261   Finally, by analysing congruences between MOTUs and dieback gradients using IndVal

262   analyses, we highlighted species-specific responses to forest dieback. We significantly

263   associated MOTUs to low and medium dieback level categories but not to high dieback level

264   one (Supplementary Table III). These MOTUs could be linked to specific forest dieback

265   conditions with particular environmental niches and therefore be considered losing over the

266   general dieback gradient. The remaining species may either be too poorly sampled across plots

267   to assess significant linkage or likely to be spread across multiple levels of dieback.

268

269   *Environmental drivers of species turnover*

270   When assessing the contribution of our eight variables as drivers of the compositional

271   turnover across zeta orders, we found that distance between plots played an important role in

272   explaining the observed variance, especially in a two-plots to 10-plots comparison—hence

273   dominated by rare species—and even greater at zeta orders above 28, thus on common species

274   (Fig. 8A). This increase in geographic effect was in accordance with the relative number of

275   plots in each respective district, for which communities were significantly different (Fig. 4B;

276   Supplementary Table II). Besides distance, both altitude and canopy openness, and to a lesser

277   extent density of large trees played a major role in driving community composition of rare

278   species (Fig. 8A, zeta order 2). Similar results were observed at zeta order 10, but TreM

279   diversity and volume of deadwood became more important, while canopy openness was less

280   impactful. Finally, the bigger the zeta order was, the more impact had TreM diversity and

281   deadwood volume on-site, as only these two environmental factors (distance excluded) were

282   found to drive compositional turnover (Fig. 8A, zeta order 50). Interestingly, TreMs density did

283   not affect overall compositional turnover (data not shown), suggesting that TreM diversity

9

284   rather than TreM amount drove community composition. The eight tested variables explained

285   20 to 40% of the species composition turnover variance across all zeta orders. Distance

286   excluded, the seven remaining variables only accounted for 15 to 20% of the variation, hence

287   most of the variance remained unexplained (Fig. 8B). This suggests a complex multifactorial

288   effect of environmental factors—but not random assemblies (Fig. 3C–D; Supplementary Fig.

289   5C–D)—, with many variables yet to be explored, in driving the insect community composition

290   turnover of both rare and common species.

Discussion

Marked declines in insect abundance, biomass and species richness have recently been quantified in Europe[2]. The causes of insect decline are multifactorial with rapid climate change identified as one of the major drivers. Here we investigated whether forest disturbances such as drought-induced forest decline and subsequent salvage logging have an impact on flying insects. Surprisingly, insect richness remained stable regardless of the extent of dieback or logging. However, species composition changed significantly with the level of dieback, driven mainly by rare species. Comparable changes in insect species composition but stable insect species richness have been observed in response to pest-induced forest dieback of other tree species[22,33], and the fact that similar observations have been made in aquatic insect communities over time[17] suggests that our findings are indicative of broad patterns.

While species richness is similar, we highlighted that forest dieback caused significant changes in insect community compositions for the whole dataset. In an increasing impact order of dieback level, Lepidoptera, Hemiptera, Diptera and Hymenoptera all showed changes in community composition across the dieback gradient. Furthermore, all functional groups studied (*i.e.* floricolous / non-floricolous and parasitoids / non-parasitoids) showed compositional turnover driven by forest dieback. In line with previous study, we demonstrate how widely impacted biodiversity can be from dieback regarding both taxonomic and functional assemblages[34]. Only Coleoptera have had no detected effect of disturbances in general on their communities. This is surprising as Coleoptera have been shown to be significantly affected by both forest diebacks and salvage logging[35]. But while they were the sole group bearing significant change in species richness across districts, this absence of community change across dieback gradient had been already observed on saproxylic beetles in the same sampling area[36] but may also be an artefact of the use of Malaise traps and a biased or lower sampling efficiency of Coleoptera in Sault plateau.

Salvage logging practices had no impact on species richness nor on community composition in the Pyrenees. This result has to be taken with caution since the salvage logging intensity and the associated amount of deadwood removed from our plots might have been too small to significantly impact insect fauna. Indeed, as the slopes of the sampled mountain forest plots were steep and logging mechanically performed using skidders and cables, we suppose that both reasoned management practices and the difficulty to heavily harvest deadwood mitigated salvage logging impacts on insect biodiversity. In addition, winning and losing taxa from salvage logging[13] may have had overall counter-balancing effects. Indeed, we found a

325 weak but expected negative effect of salvage logging on the prevalence of Coleoptera similar

326 to other studies[13]. However, we think the first explanation of low salvage logging intensity is

327 more plausible and supported by the fact that we did not detect any impact of salvage logging

328 on deadwood amount in our study plots[36]. Hence, one must remain cautious with the lack of

329 detectable impact of salvage logging in our study. Furthermore, as sampling was happening in

330 non-old-growth forests, this may have left a sufficient amount of coarse woody debris on-site

331 to prevent severe impacts on insects. Insect communities in managed stands such as those we

332 studied may be poorer and more homogenous than in more mature forests, reducing our ability

333 to detect changes in species richness. Additionally, our sampling was carried out more than 10

334 years after the onset of climate-induced tree dieback and subsequent logging, giving plenty of

335 time for a recovery of species richness, even though extinction debt (*i.e.* a time-delayed negative

336 response of a taxon to an environmental disturbance) has already been observed in Diptera after

337 29 years in *Quercus* spp. dominated forests of Southern France[37].

338 Overall, our findings underline that forest dieback favouring insect biodiversity[14] may

339 be based on the response of well studied groups, *i.e.* "those insects that we love and cherish"[38].

340 Indeed, rather than an increase in biodiversity, we demonstrated that species richness remained

341 steady regardless of dieback level. Even though they did find an overall positive response,

342 Moretti *et al.*[15] highlighted by breaking down the datasets at lower taxonomic level that insect

343 diversity could respond both negatively and positively to fire-induced forest diebacks. They

344 found that the winners were among the most studied groups of insects (*i.e.* hoverflies, bees,

345 social wasps and ground beetles) while only one losing insect group was identified (weevils)[15].

346 This implies that other, unidentified groups were potentially "losers" but less likely to be

347 reported. Similar bias toward these well-known insect groups is also identifiable in studies on

348 pest-induced forest dieback with an additional focus on red-listed species[10], or in the studies

349 considered in Thom and Seidl's meta-analysis[14]. From this detailed taxonomic view[15] and

350 recent comprehensive studies[22,33], insect positive response to disturbances must be taken

351 cautiously as it could be biased towards the response of the better studied insect groups only.

352 Previously reported global patterns may thus not be fully representative. Here, by including

353 hyper-diverse and understudied groups ("dark taxa")[32], the overall stable species richness that

354 we found across both the dieback gradient and stand types highlights a nuanced and

355 idiosyncratic response of insect orders and functional groups such as floricolous / non-

356 floricolous and parasitoid / non-parasitoids insects to forest disturbances. Additionally, our

357 results in both insect species richness and compositional turnover emphasize the limitation

12

358   inherent to the sole use of species richness as a metric[39] and the need to look more closely at

359   community composition and functional changes to notice disturbance effects[40].

360   Here, we highlight that various dieback levels would rather promote different insect

361   community assemblages and be part of natural succession dynamics (*i.e.* ecological change

362   occurring in a predictable way after disturbance), with equally weighted response of species

363   either winning or losing throughout the dieback intensity gradient. Hence, management policies

364   based on species richness metrics and promoting dieback only to favour biodiversity may hold

365   potential risks by driving a catastrophic decline of insects through a potential homogenization

366   of dieback levels and associated species assemblages with more generalist species. Albeit

367   Borges and collaborators[40] highlighted the issue in the very distinct island ecosystem, their

368   concern can be extended to freshwater streams[17] or disturbed forest environments as in the

369   present case.

370

371   In our study, zeta decline fitting a power-law regression implied a compositional turnover

372   of the total dataset driven by environmental factors rather than stochastic events[29]. The zeta

373   ratio analyses showed that species are indeed retained differently within each dieback category.

374   Thus, the lack of core set of species observed for Lepidoptera may be attributed to a relatively

375   low moth diversity associated with silver fir-dominated forests with community composition

376   mainly shaped by turnover processes[41]. On the other hand, sampling efficiency could explain

377   the absence of common species retained for Hemiptera for the different dieback categories or

378   stand types tested[29]. Nevertheless, the prevalence (*i.e.* the presence probability of a taxa within

379   all the plots grouped in a given environmental category) of these two orders remained positively

380   correlated with dieback in general, in accordance with previous observation in which open

381   canopy deriving from diebacks favours heliophilous and flower-visiting insects[15]. Interestingly,

382   the prevalence of Coleoptera was favoured by low dieback level while species richness across

383   the dieback gradient was similar and the species retention rate decreased. Even though we could

384   not finely assess feeding guilds, these observations may support high turnover processes and

385   competition deriving from specialist species' population decrease as factors shaping Coleoptera

386   communities[9]. However, salvage logging had a slight but significant negative impact on

387   Coleoptera prevalence, in line with previous observations[13,36,42]. In addition, each dieback

388   gradient hosted a specific core set of one or more common species composed of Diptera and

389   Hymenoptera at both low and medium dieback levels. The reduced openness associated with

390   healthy stands—hence low dieback level—might also support the reduced prevalence of both

391   Diptera and Hymenoptera observed in these environments, as most of the species recovered for

13

392 these orders are considered floricolous. Dieback might also increase environment complexity

393 and support the positive effect of medium dieback level that we observed on Hymenoptera,

394 especially parasitoid ones[23], rather than a negative bottom-up effect of the host condition on

395 parasitoid species[24].

396 A compelling observation in terms of functional diversity is that while both Diptera and

397 Hymenoptera displayed strong compositional turnover at high dieback level, their drastic

398 changes in community composition drove the turnover observed for floricolous and parasitoids

399 functional assemblages. It is known that both floricolous and parasitic Hymenoptera can indeed

400 be significantly impacted by structural changes in silver fir-dominated mountain forests[43].

401 However, our results emphasize the ecological importance of some species-rich yet often

402 overlooked taxa on wide functional groups. Furthermore, it highlights the need for more work

403 to investigate whether climate-induced forest dieback can deeply impact functional efficiency

404 and ecosystem services through these taxa[3].

405 In a more detailed view, within each studied order or functional group, the winner and

406 loser taxa were relatively rarely occurring species in our dataset. This pattern is comparable to

407 previously reported taxonomic dissimilarities driven by rare species[34] and consistent with

408 models showing that selective perturbations directed towards rare species generate more

409 dynamical effects on the ecosystem[44]. As rare species have a positive impact on ecosystem

410 multifunctionality[45,46], these rare taxa have high patrimonial value and thus their importance in

411 forest ecosystems functioning should be further investigated[34].

412 Geographic distance—regardless of the zeta order considered—played a significant role

413 as a macroecological factor shaping the different insect communities of the two districts and,

414 as for the aforementioned winner and loser taxa, mostly taxa with low prevalence shaped

415 compositional specificities of each district. As expected, many forest features tested (*i.e.*

416 altitude, density of large trees, canopy openness, total volume of deadwood and TreM) also had

417 a significant effect in shaping insect community composition, indicating a complex response to

418 disturbances. Remarkably though, environmental factors explaining most of the compositional

419 variance differed according to the observed zeta ratio. At low scale (zeta order = 2), community

420 changes were mainly driven by canopy openness, density of large trees and altitude. This result

421 support the role of distance and stand structure heterogeneity in favour of biodiversity[47] and is

422 in line with previous observations on these features affecting the rare species composition at

423 low zeta order[22]. However, these environmental features at large scale (zeta order > 10) were

424 no longer the main drivers of biodiversity changes. Instead, both TreM diversity and the amount

425 of deadwood were the main factors influencing community composition when considering the

whole dataset. Interestingly, TreM diversity rather than density had most impact on insect diversity (data not shown)[48], similarly to deadwood diversity and amount in other studies[47,49]. However, as TreM diversity is linked to TreM density, it therefore might indirectly and at different intensity also influence biodiversity[48]. Furthermore, and contrary to previous report[47], we found deadwood amount to be as valuable as stand structure for biodiversity conservation, depending on the geographic scale managed[42]. This result further supports the general call on acknowledging and considering the importance of deadwood to biodiversity[42]. Interestingly, a previous study successfully linked hoverfly diversity to plot connectivity and the combined effect of environmental factors from both stand and landscape scales[37]. The shift in environmental variables driving community assemblages according to the zeta order considered we highlighted here support such geographic scale effect and help disentangling the key drivers in action. This may potentially be extended to most aerial insects with higher dispersal ability, which are the main representatives of our Malaise-trapped dataset. Hence, while species-specific response to environmental factors must be accounted for, geographic scale of the conservation area, especially through stand and landscape scales[37], as well as dieback intensity should also be considered when managing stand heterogeneity and environmental factors to favour forest insect biodiversity[47].

Finally, our study confirms DNA metabarcoding of Malaise trapped samples as an efficient approach for biomonitoring changes in species-rich insect communities that include "dark taxa" which collectively represent the bulk of forest insect diversity[22]. Indeed, our metabarcoding pipeline coupled with recent efforts to complete the DNA barcoding reference library of European insect fauna[19,50] allowed the inventory of nearly 3000 insect species, more than half of which were identified to species level, including hyper-diverse taxa such as Diptera and Hymenoptera[51]. These two taxonomic groups were the most diverse in our study, like in other Malaise trap environmental surveys[52,53]. Although Malaise-trapped samples are biased toward flying insects[54], the zeta decline analyses indicated that our dataset incorporated rare species as non-stochastic events and tracked changes in both species richness and community composition at both large and small spatial scales while overcoming sampling district discrepancies[29]. Scaling up our approach to national and continental levels would help to monitor insect biodiversity and potential decline, to provide an understanding of the environmental drivers of biodiversity loss in a rapidly changing climate, and allow for regular assessment of the efficiency of conservation and forest management policies.

15

Methods

Study sites, stand description and insect sampling

Fifty-six plots were set up in 2017 in silver fir forests in Central (Aure valley) and Eastern (Sault plateau) French Pyrenees (see "*Supplementary Information – Plot list*").

Tree dieback was assessed using the ARCHI method[55]. The 20 closest silver firs from the plot center (defined by the Malaise trap), belonging to the dominant layer and with a diameter at breast height (dbh) > 17.5 cm, were examined with binoculars. The proportion of affected trees (*i.e.* trees expressing signs of climate-induced dieback and comprising stressed + resilient trees) and dying trees was used as the proxy of the decline at the stand level to define a three-category gradient: 18 low dieback level plots (no dying trees and ≤ 60% of affected trees), 23 high dieback level plots (at least 1 dying tree and > 60% of affected trees) and 15 medium dieback level plots (no dying trees and > 60% of affected trees).

For each living or dead tree belonging to a fixed-angle plot (ratio 1/50), we recorded its status (living, snag, log), tree-species, dbh (minimum dbh to be recorded: 17.5 cm for living trees and logs, 7.5 for snags). Tree-related microhabitats (TreMs) were recorded on living trees and snags according to the Larrieu and collaborators' typology (47 types)[48]. TreM diversity corresponds to the number of TreM types recorded per plot. Very large trees were defined as living trees above 67.5 cm in dbh. Since we used the ratio 1/50 to fix the angle, basal area matches with the number of trees observed. Each deadwood item (length > 1 m) was measured in length and diameter to calculate its volume. Mean plot area was about 0.3 ha, depending on the dbh and location of the largest trees. Additionally, canopy openness was evaluated by using a spherical densiometer[56] at the plot centre and 10 m from the centre in four crossed directions (thus 5 locations per plot). Canopy cover was calculated for each location as the proportion of the 68 points of the densiometer screen that was intersected by cover. Plot value resulted as the mean of the five counting.

Stand type was described on an area of 1 ha using the Index of Biodiversity Potential (IBP)[57]. The IBP is a biodiversity evaluation tool (according to Larsson[58]) combining ten historical, structural and compositional key factors for forest-dwelling species which are easily and directly measurable in the field[59]. The measurement of these factors provides a rapid assessment of the theoretical biodiversity hosting capacity of the stand and facilitate the identification of the key environmental variables manageable for better biodiversity potential. In our study, the main factors accounting for healthy, disturbed and salvaged conditions were stand structure (canopy openness), tree health (deadwood volume and number of living trees)

16

493 and the awareness of management activities influencing deadwood volume and density of

494 habitat-trees available on-site.

495   Insect sampling was conducted over late spring–early autumn season from May 15 to

496 September 15, 2017 using 56 Townes-style Malaise traps with black walls and white roof[53].

497 One Malaise trap was placed at the centre of each of the 56 one-ha plots. Malaise trap samples

498 were filled with a mixture of 20% mono-propylene glycol and 80% pure ethanol. Samples were

499 retrieved once a month, giving a total of four samples per plot over the sampling period (124

500 trap days), for a total of 224 samples. After collection, all samples were stored at 4°C until

501 laboratory processing.

502

503 *Laboratory processing and DNA extraction*

504   Insects were first filtered from the trapping solution and rinsed with ultrapure Milli-Q

505 water to remove mono-propylene glycol residue. Insects were then placed within sterile and

506 disposable Petri dishes on clean absorbing paper to dry overnight at ambient temperature.

507   Once dried, insects were size-sorted using decontaminated forceps. Insects larger than

508 a European honey bee (*Apis mellifera* Linnaeus, 1758) were removed and only the head or a

509 part of the abdomen was retained in order to reduce the biomass and improve the detection of

510 rare or small species[60]. Insect bulk samples were then ground and homogenized into fine

511 powder using disposable BMT-50-S-M gamma sterile tubes (IKA) with 10 steel beads with an

512 Ultra Turrax Tube Drive grinder (IKA). Ground bulk samples were then conserved at -21°C.

513   DNA extraction of 25 mg (±2 mg) of insect powder was performed on silica columns

514 using a standard DNeasy® Blood & Tissue extraction kit (QIAGEN) (see:

515 www.qiagen.com/handbooks for further information). For each ground sample, we took 25 mg

516 (±2 mg) of insect powder in a 1.5-mL microcentrifuge tube using spatula previously

517 decontaminated with 4% Decon® 90 solution and autoclaved. During powder sampling, empty

518 1.5-mL microcentrifuge tubes containing 200 µL ATL buffer were left open and changed every

519 nine samples to control for potential cross-contamination with volatile insect powder and all

520 were processed as extraction controls (EC). Mock communities of 248 Asian insects were used

521 as positive controls. Negative (NC), extraction (EC) and positive (PC) controls were processed

522 down to sequencing. Lysis was performed with horizontal shaking overnight at 56°C in 180

523 µL ATL buffer and 20 µL proteinase K. All vortex steps were replaced by handshaking to

524 reduce DNA degradation and DNA was eluted into 80 µL of AE buffer following 15 min

525 incubation on the silica column at ambient temperature, and a second elution with the previous

526 eluate after 5 min incubation on the silica column. Each sample was quantified using Qubit®

2.0 fluorometer dsDNA High Sensitivity kit (Invitrogen) and eluate subsamples were diluted to 2 ng/µL. Both stock solutions and diluted eluates were stored at -21°C.

A twin-tagging dual-indexing approach was used for the sequencing library preparation[61] (see "*Supplementary Information – Primer list*"). We targeted a 313-base pair (bp) fragment of the mitochondrial DNA Cytochrome oxidase *c* subunit I (COI) gene. PCR amplification was done using the mlCOIintF forward primer 5'–GGWACWGGWTGAACWGTWTAYCCYCC–3' and jgHCO2198 reverse primer 5'–TAIACYTCIGGRTGICCRAARAAYCA–3'[62,63]. To facilitate twin-tag multiplexing, each of the 96 primer pairs was synthesized with a 7-bp tag at their 5' end, differing by at least three bp and avoiding G nor TT at the 3' end to prevent a succession of the three same nucleotides once attached to the primers. Furthermore, A, T, CA, CG, GC and GT nucleotides were added to the tag to act as heterogeneity spacers[64] to increase complexity and shift the reading frame for better sequencing quality on Illumina MiSeq platform (see "*Supplementary Information – Primer list*" for a complete list of designed and used primers). Shifting was in accordance with green/red light balance on v2/v3 Illumina MySeq technology[65]. Neither proof-reading nor hot-start Taq polymerases were used for PCR amplification due to Inosin bases in the reverse primer[66].

Prior to amplicon PCR amplification, we tested the optimal number of cycles and DNA quantity added to the PCR-mix by quantitative PCR (qPCR) using a LightCycler® 96 Instrument (Roche) with the KAPA Library Quantification Kit (Illumina). The qPCR conditions were: one preincubation step of 95°C for 5 min, followed by 40 cycles of 95°C for 30 sec, 45°C for 60 sec and 72°C for 90 sec, with a final extension at 72°C for 10 min and a high-resolution melting step of 95°C for 60 sec, 40°C for 60 sec, and an acquisition gradient from 65°C to 97°C. The qPCR reactions with a final volume of 15 µL were prepared with 7.5 µL of KAPA qPCR mix (2X), 0.3 µL of each forward and reverse primers (10 mM), 3.9 µL of extra pure molecular grade water and dilution series of 3 µL DNA template (set at different concentrations: 8, 4, 2, 0.4 and 0.2 ng/µL) were performed to investigate inhibitions. The optimal number of PCR cycles was designated by where the exponential phase transitioned to plateau phase. PCR amplifications were thus run for 23 cycles under the same conditions as the qPCR, in 25 µL reaction volume as follows: 2.5 µL of buffer green (10X), 0.5 µL of dNTPs (10 mM), 1 µL of each forward and reverse primers (10 mM), 0.2 µL of Metabion mi-Taq (5 U/µL), 16.8 µL of extra pure molecular grade water and 3 µL of DNA template at 2 ng/µL. Each sample was PCR-amplified three times in different 96-well plates with three different

18

561 twin-tag sequences, to allow independent tracking of the three PCR products after

562 amplification. PCR products were purified using CleanPCR magnetic beads (CleanNA) at a

563 ratio of 0.8/1 µL of PCR product. Two steps of rinsing with 200 µL of 70% ethanol were

564 performed before final elution in 25 µL TE buffer (1X). Purified PCR product was then

565 quantified in duplicates on a FLUOstar OPTIMA microplate reader (BMG Labtech) with

566 Quant-iT^{TM} PicoGreen® dsDNA assay kit (Thermofisher) following manufacturer's protocol.

567 Equimolar pooling of the samples was carried out for each plate with a total DNA quantity of

568 1 µg of purified amplicon for a final volume of 60 µL. A second indexing for all PCR plates

569 was conducted using the TruSeq DNA PCR-free Library Prep kit for PCR-free ligation

570 (Illumina) with 18 different indices (one per 96-well plate) following the protocol given in

571 Leray *et al.* (2016)[61]. Library preparation and sequencing of the total 224 was done on five

572 MiSeq runs – v3 (Illumina) with 600 cycles.

573

574 *Use of positive controls and blanks in demultiplexing*

575 Mock communities of known species composition with DNA barcode sequences

576 available were used as PCs. The 248 species used in the mock community all have distributions

577 restricted to east Asia, and hence allow for the examination of tag-jumping in the Illumina

578 sequencing process. PCs were sequenced and bioinformatically treated the same way as bulk

579 samples to ensure comparability. We focused on a restrictive approach, favouring false

580 negatives to false positives (See Alberdi *et al.,* 2018[67] for more information on bioinformatic

581 set-up). To decide both on the number of PCR replicates in which MOTUs must appear and

582 minimum read numbers per PCR, we used BLAST+[68] on the PCs to find the set up in which

583 we could recover the most MOTUs with 100% match and no MOTUs below 97%. Both NC

584 and EC were checked for reads after demultiplexing to ensure that there was no cross-

585 contamination and removes potential MOTUs recovered in controls from the total dataset.

586

587

588 *Reads demultiplexing and taxonomic assignment*

589 AdapterRemoval *ver. 2.2.2*[69] was used to trim the twin-tagged adaptors of the reads and

590 we employed sickle *ver. 1.33*[70] to perform paired-end quality trimming. Error correction using

591 Bayes Hammer was done via SPAdes[71] *ver. 3.12.0,* followed by paired-end merging with

592 PandaSeq[72] *ver. 2.11*. PCR replicates were treated following DAMe pipeline[73]. Heatmaps of

593 primer tag combinations versus read numbers were generated using R *ver. 3.6.1*[74] to control for

594 indexing quality. Our twin-tagging approach allowed us to easily identify errors and problems

595 at this step. Renkonen Similarity Indices (RSI value[75]; were calculated for unique sequences

596 found in at least 3 PCR replicates and represented by at least 3 reads (See section 'Use of

597 positive controls and blanks in demultiplexing' for arbitration of required numbers of PCR

598 replicates and reads). Average read size was plotted in R to ensure filtering and trimming

599 quality, and VSEARCH[76] *ver. 2.8.1* was run to remove chimeras.

600   Sequences were clustered into Molecular Operational Taxonomic Units (MOTUs) with

601 a 97% similarity threshold[77] using SUMACLUST *ver. 1.0.20*[78]. Clustering quality was

602 controlled via R, using the LULU approach[79].

603   Finally, taxonomic assignment was performed against BOLD System database[27] in

604 April 2019 using a 97% threshold similarity[22] with bold R-package *ver. 0.9.0*[80]. For each

605 MOTU, only unambiguous taxonomic assignments were kept, meaning that when multiple

606 names at a given taxonomic rank were attributed, only common and unique taxonomic

607 assignments at higher taxonomic levels were kept. In the case of multiple MOTUs with an

608 identical and unambiguous species name assignment, all related MOTUs and their respective

609 presence in sampled plots merged into a single species-specific MOTU. Functional traits were

610 associated manually to MOTUs according to family level whenever possible (see

611 "*Supplementary Information – Family functional traits*" for more details). Parasitoid and

612 floricolous families were mainly assessed from[81, 82, 83] and with the help of expert taxonomists.

613

614

615 *Statistical analyses*

616   All statistical analyses were carried out in R *ver. 3.6.1*[74]. Read-sequencing data were

617 transformed in incidence-based data to account for presence / absence of MOTUs given that

618 abundance data from metabarcoding can be misleading for metazoans[84].

619   We generated accumulation curves using *iNEXT* package *ver. 2.0.20*[85] with incidence-

620 based frequency dataset parameters and Hill number q value set to 0 to reflect truly observed

621 MOTU diversity, regarded as species proxy. Total number of MOTUs was plotted against both

622 sampling units and sample coverage for the late spring–early autumn sampling period as well

623 as for each four-weeks time series (named May, June, July and August from their respective

624 starting months). MOTUs belonging to the five most diverse taxonomic orders represented in

625 our dataset (*i.e.* Coleoptera, Diptera, Hemiptera, Hymenoptera and Lepidoptera) as well as for

626 the remaining 10 Orders (*i.e.* Blattodea, Ephemeroptera, Mecoptera, Neuroptera, Orthoptera,

627 Plecoptera, Psocodea, Raphidioptera, Thysanoptera and Trichoptera) pooled into an "Others"

628 category, were also plotted against sampling units to account for respective recovery success.

629   We also used both Chao and first order Jackknife (Jack1) methods to have different biodiversity

630   estimates from the 'specpool' function of R package *vegan ver. 2.5-6*[86]. These estimators also

631   extrapolate species richness within the sampled area in regards to sampling effort and based on

632   incidence data.

633   Impacts of tree diebacks on species richness was assessed using generalised linear

634   model (GLM). For the total dataset, each taxonomic order category of the five main

635   representatives (*i.e.* Coleoptera, Diptera, Hemiptera, Hymenoptera and Lepidoptera) or over the

636   four ecological groups (*i.e.* floricolous / non floricolous adults and parasitoid / non-parasitoid

637   larval feeding guild), GLM was first tested with both Poisson and quasi-Poisson distribution,

638   the latter always retained for its better fit. Then, each model tested the effect of the dieback

639   gradient, stand types, their interaction as well as the effect of districts. When no interaction was

640   found, a simpler additive model was tested. In all nine studied cases, quasi-Poisson distribution

641   and additive model characteristics were retained, as follow: `model <- glm(richness ~`

642   `dieback + salvage + district, family = quasipoisson)`. We finally

643   applied type II Anova from *car* package *ver. 3.0-8*[87] on the model to test the significance of

644   each parameter on the species richness of the studied category.

645   To compare insect community composition across disturbances gradients (*i.e.* tree

646   dieback intensity and stand types) and districts, we performed GLM on a generated mvabund

647   object using *mvabund ver. 4.3.1*[88] as follow: `Mod_Tot <- manyglm(Tot_mvabund ~`

648   `X$Dieback + X$Type + X$District, family="binomial")`. Reported values

649   are the two extremes resulting from a 10 runs range. Post-hoc Holm correction was then applied

650   on the 10 p-values recovered to assess for the significance of potential compositional changes

651   in community compositions. Ordinations for these GLM were performed using *ecoCopula ver.*

652   *1.0.1*[89].

653   Prevalence of the five most represented taxonomic orders in regards to dieback levels

654   and stand types was assessed by a "fourth-corner model" analysis using 'traitglm' function from

655   *mvabund ver. 4.3.1*[88] R package. For each insect MOTU, taxonomic order was defined as a trait

656   and GLM analyses run to define traits significantly associated to environmental covariates.

657   To determine potential indicator species for each dieback gradient, we used IndVal

658   analysis from 'multipatt' function of *indicspecies* R package *ver. 1.7.9*[90]. Again, reported p-

659   value ranges are resulting from 10 independent IndVal runs with post-hoc Holm corrections

660   subsequently applied.

661   We also assessed the importance of rare species in species assemblages specific to both

662   the gradient of dieback intensities and the sampling districts. Here, we defined a rare species

from its frequency of appearance in a given dieback category or district, with occurrence (*i.e.* number of plots of the relative category in which we found the species) as an index of relative abundance. Hence, MOTUs more frequently sampled throughout the plots were defined as more abundant and more common. We then generated heatmaps of relative abundance using the phyloseq *ver. 1.30.0*[91] package for R. As rare species—defined from the multiple dimensions of rarity[92]—can therefore be virtually locally abundant in a given dieback category or district by the use of occurrence only to define our MOTU rarity, relative abundance used in heatmaps was generated from 100 randomized bootstraps of the total dataset to virtually increase sampling effort. Bootstrapped plots were then concatenated to recover a bootstrapped score giving the relative abundance. From this score, we could infer rarity in terms of habitat specificity, with common species occurring across spatial scale and/or dieback gradient, and rare species occurring only in a few plots in a particular district or level of dieback. Each of the three dieback categories (*i.e.* low, medium and high dieback levels) was rescaled to account for the difference in sample representativeness, as such that each category was depicting one third of the total dataset. Each MOTU relative abundance was then calculated in percentage from the bootstrapped score for these rescaled categories, so that a MOTU present in all plots of one particular dieback category was present at 33% in the respective category, for a total of 100% across the total dataset if appearing everywhere. Regarding abundance per district (*i.e.* Aure valley and Sault plateau, as well as total dataset), MOTUs' relative abundances were calculated in percentage based on the bootstrapped score with no further rescaling per district.

Species compositional changes along the total sampling area was assessed by calculating the number of species shared across all 56 experimental plots using zeta diversity (*i.e.* extension of beta diversity over *i* zeta order (ζ_i) or plots)[28] using zetadiv R package *ver. 1.2.0*[30]. Two models were computed to perform plot comparisons either following a nearest-neighbour (NN) scheme or considering every comparison possible (ALL scheme) to assess for heterogeneity driven by distance. The two models (NN and ALL) were also fitted both to power-law (*pl*) and exponential (*exp*) regressions to define whether the signal in our data set was driven more by ecological niche and rare species turnover rather than by stochasticity, respectively[29]. To assess the better fit to a model between *pl* and *exp*, we chose the lowest score from Akaike Information Criterion (AIC) which imposes penalty to each model in regards to its number of parameters, and thus favour models satisfying parsimony criterion[93]. For both dieback gradient (*i.e.* low, medium and high dieback categories) and stand types (*i.e.* healthy, disturbed and salvaged), zeta ratio (*i.e.* the species retention rate, giving the probability of a common species to be retained in the community across zeta order here referring to plots) was

697 generated for each of the five most represented taxonomic orders and the four different

698 functional groups identified, as well as for the mean zeta ratio of the respective dieback category

699 or stand type on the total dataset.

700 Estimations of environmental variables contribution to insect species assemblage changes

701 were performed with zeta multi-site generalized dissimilarity modelling (zeta.msgdm) in

702 zetadiv R package *ver. 1.2.0*[30] at different zeta order, ranging from two to 50. In total, eight

703 environmental variables were tested with zeta.msgdm: geographic distance to the nearest plot,

704 altitude, canopy openness, total amount of deadwood in m^3 per ha, the basal area per ha, the

705 tree diversity per ha, the density of very large trees ($\varnothing > 67.5$ cm) per ha and the TreM diversity

706 per ha. zeta.msgdm was calculated using I-spline models[94].

Acknowledgements

Data availability

All scripts and datasets used for analyses are publicly available at the following GitHub repository: https://github.com/Lucasire/Malaise_FR_2017. Raw sequencing data will be available on NCBI upon publication with the following accession number: PRJNA702908.

Author contributions

This study was conceptualized and designed by CLV, EAH and CB. Forest plot selection and sampling were designed by LL and CB. Sample processing, wet-lab experiments and sequencing were performed by LS, PSY with the help of AB and BC. Bioinformatic demultiplexing was done by LS and PSY with the help of DWY and CW. Ecological and environmental analyses were conducted by LS and WC, with the help of JC, CB, DWY and DF. LS led the writing of the manuscript. All authors contributed substantially to the interpretation and discussion of the results as well as to the revision of the article. All authors approved the submitted version.

Conflicts of interest

The authors of the manuscript declare no competing interest.

Figures and Tables

Fig. 1: **iNext species accumulation curves for Malaise trap samples and temporal turnover.**
(**A**) Species richness (MOTUs) per sampling unit recovered through metabarcoding of Malaise trap samples (solid lines). Incidence frequency-based species richness extrapolations (dashed lines) were performed over the total sampling period (from May 15 to September 15, 2017) (green curve) as well as for the following four-weeks temporal sampling categories: May, June, July, August to assess the potential variations in both sampling effort and species richness (pink, yellow, cyan and violet curves, respectively). Sampling units are the total number of Malaise trap samples available per temporal category. Shaded area represents 95% confidence interval. (**B**) Sample coverage (sampling efficiency of sampling units to recover the expected biodiversity) per number of sampling units; (**C**) Species diversity (MOTUs) per sample coverage. Accumulation curves indicated that the 224 sampling units used in the present study allowed a 90% sampling coverage of the given area and sampling period and allowed us to recover nearly 75% of the total Malaise-trappable diversity. Estimations gave around 4000 trappable species and the need of more than 400 sampling units on-site to recover them. (**D**) Venn diagram showing the species (MOTUs) turnover in the French Pyrenees across the four-month sampling period of the study. Total number of MOTUs found every four weeks is given under the name of each starting month. Number of MOTUs shared between months is given within the diagrams. The high MOTU numbers specific to each month and the comparatively low number of shared MOTUs throughout the entire sampling period (highlighted in red) indicated an important temporal turnover of insect species.
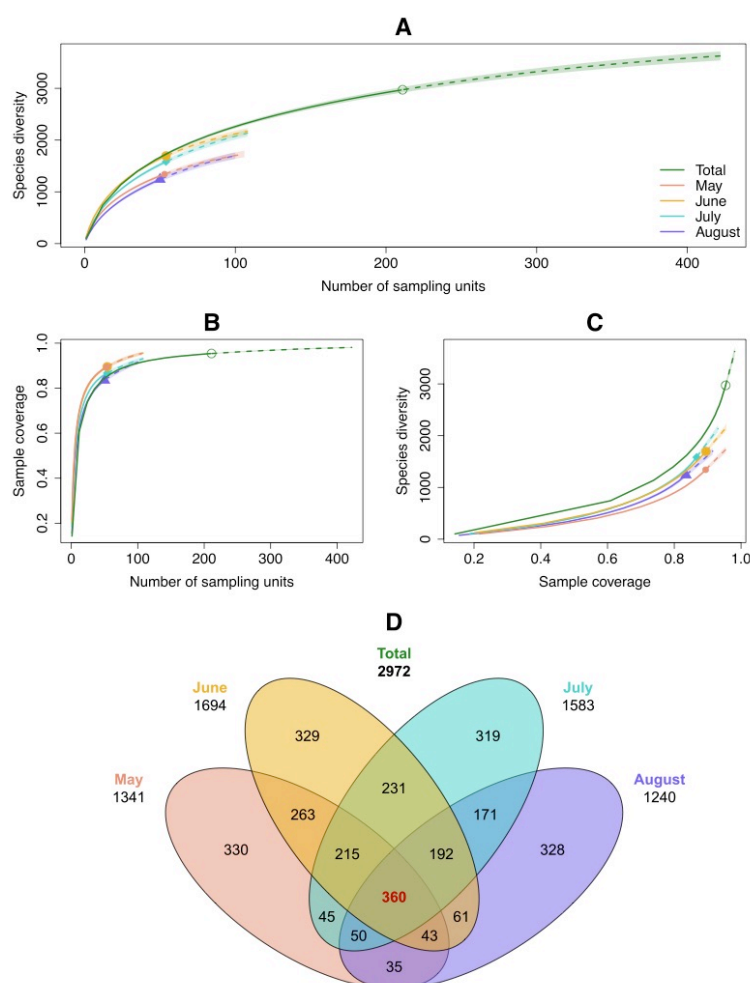
779 Fig. 2: **Taxonomic composition and Gamma diversity analysis**
780 (**A**) Pie chart of the 15 insect orders sampled showing the taxonomic affiliation and the
781 respective number of MOTUs recovered, as well as their total proportion. For clarity purposes,
782 only the five orders with highest number of species (Diptera, Hymenoptera, Lepidoptera,
783 Coleoptera and Hemiptera) are represented (in blue, red, green, brown and yellow,
784 respectively), with the 10 remaining orders (Blattodea, Ephemeroptera, Mecoptera, Neuroptera,
785 Orthoptera, Plecoptera, Psocodea, Raphidioptera, Thysanoptera and Trichoptera) clustered
786 within the "Others" category in pink. Representativeness of each insect order in the dataset was
787 consistent with the known bias of Malaise trap sampling, especially toward dipterans and
788 hymenopterans. (**B**) Incidence frequency-based accumulation curves (solid lines) with species
789 richness extrapolations (dashed lines) for each insect order. Shaded area represents 95%
790 confidence interval. As all accumulation curves were nearly plateauing, this indicated that for
791 each represented insect order, almost all Malaise-trappable diversity over the sampling period
792 of had been successfully recovered.
793

794
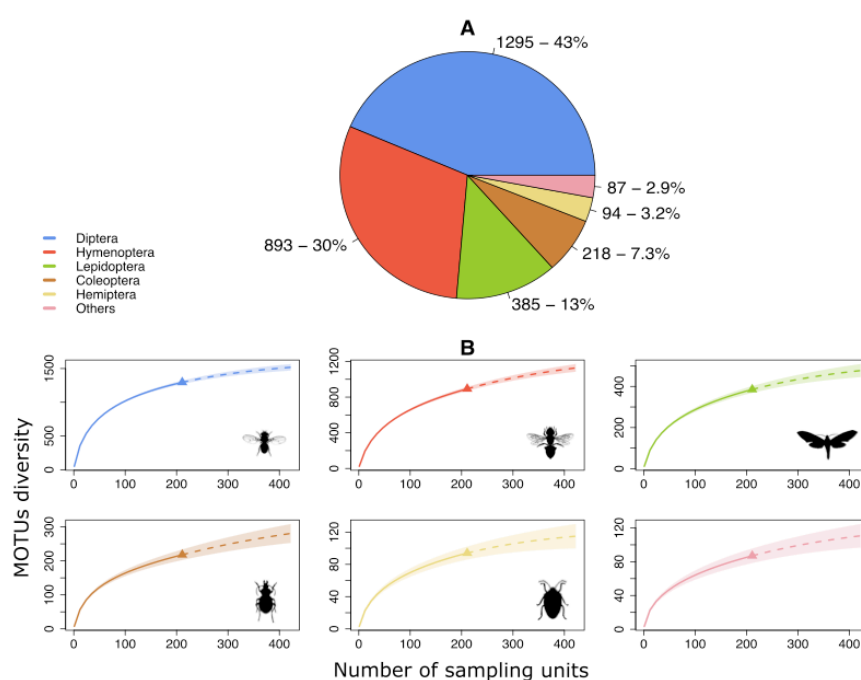
Fig. 3: **Zeta-diversity decline and model fitting on the insect fauna for nearest-neighbour combinations and assemblages.**

Zeta-diversity analyses per zeta order (*i.e.* ζ_i) here referring to plots (from $\zeta 2$ to $\zeta 56$) for two computing schemes. Representations consider model scheme that computes nearest-neighbour plot combinations and assemblages (NN). (**A, B**) Zeta-diversity decline and 0 to 1 scaled ratio of zeta-diversity decline, representing species shared and species retention rate (*i.e.* the retention probability of common species in the community) across ζ_i, respectively. Zeta-diversity decline curve (**A**) indicated high heterogeneity in insect diversity across the studied plots as fewer and fewer species were shared between 2–9 plots and almost no common species were being shared when 10 plots or more were considered. Zeta ratio (**B**) highlighted the high prevalence of the few species shared across 2–54 plots which were likely to be sampled, and the drop of species retention rate at zeta order $\zeta 28$ corresponded to geographic distance between the two sampled districts. (**C, D**) zeta-diversity model fitting to exponential and power-law regressions, respectively. Fit to exponential regression implies equal and stochastic turnover among species whereas fit to power-law regression as represented here indicates turnover driven by niche differentiation processes and rare species less likely to be found across the whole dataset.

27
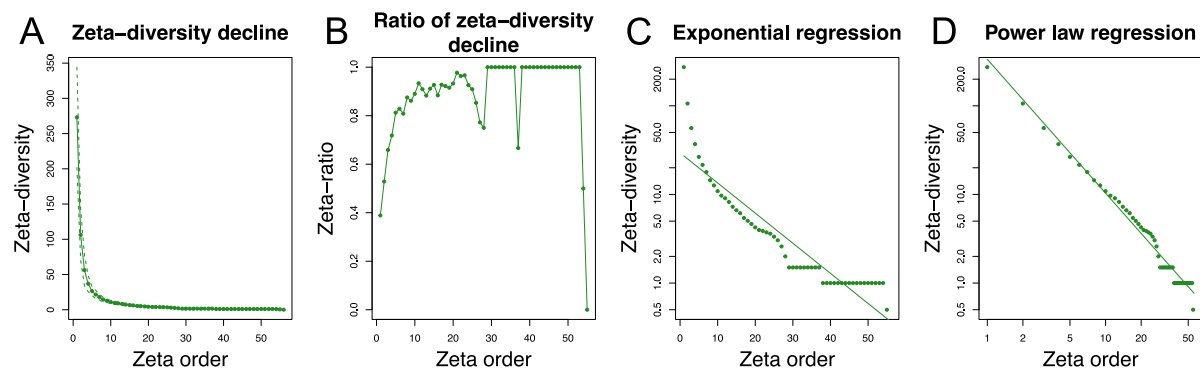
Fig. 4: **Variation in community composition across dieback gradient, stand types and geographical districts.**

Ordination plots representing variations in insect community assemblages in regards to: (**A**) dieback level conditions (*i.e.* low, medium and high) and stand types (*i.e.* healthy, disturbed and salvaged), and (**B**) to the two different geographical districts sampled. Mvabund analyses of community dissimilarities were performed using 999 bootstraps. Ordination (**A**) highlighted higher similarity in community composition between low dieback level plots than the other two dieback levels, more similar to each other but expressing high variability in community composition. No grouping of salvaged-logged plots in terms of community composition could be distinguished. Ordination (**B**) highlighted the greater dissimilarity in terms of community composition within plots of Aure valley compared with Sault plateau that hosted species assemblages more similar to each other.
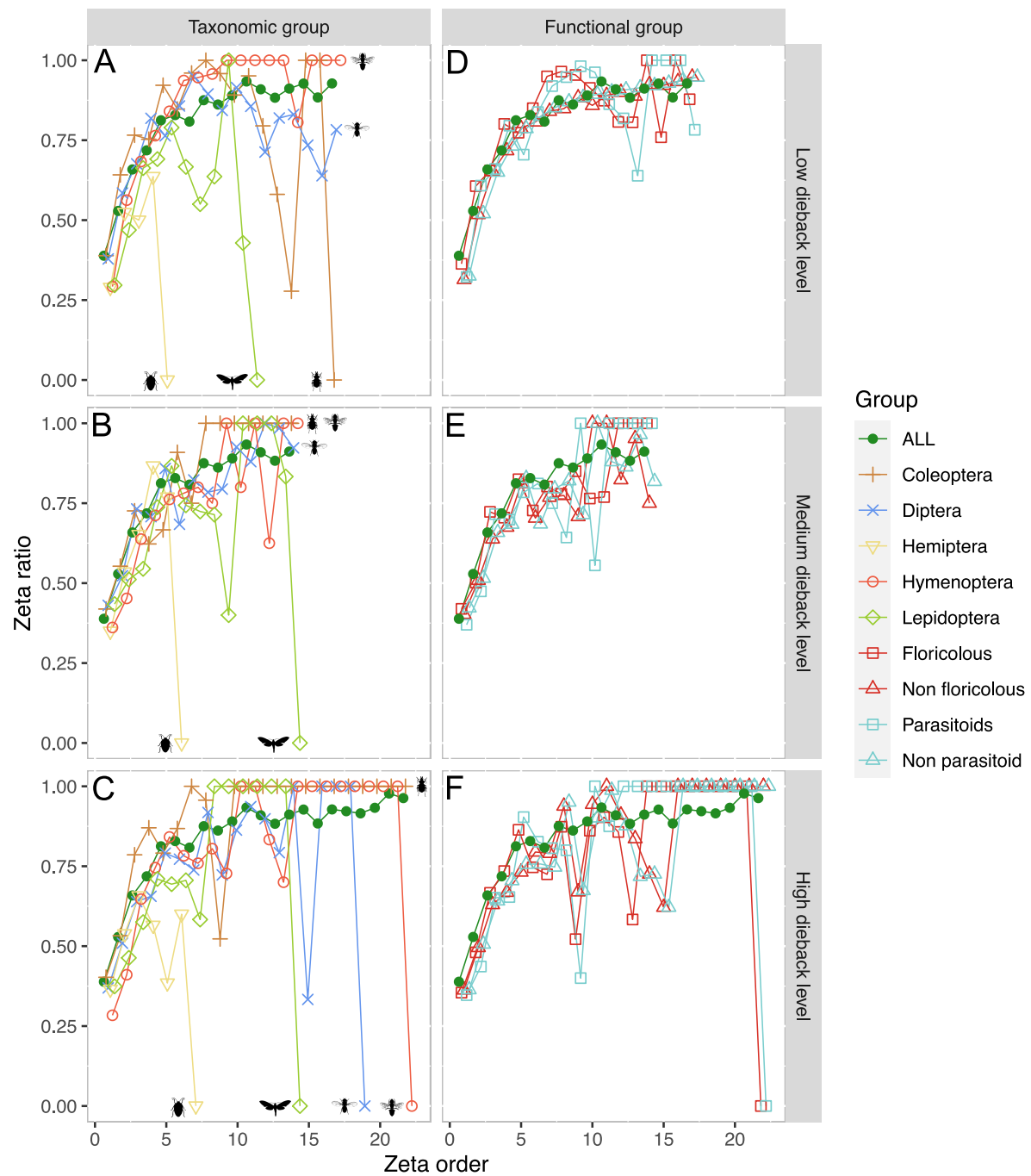
828

29
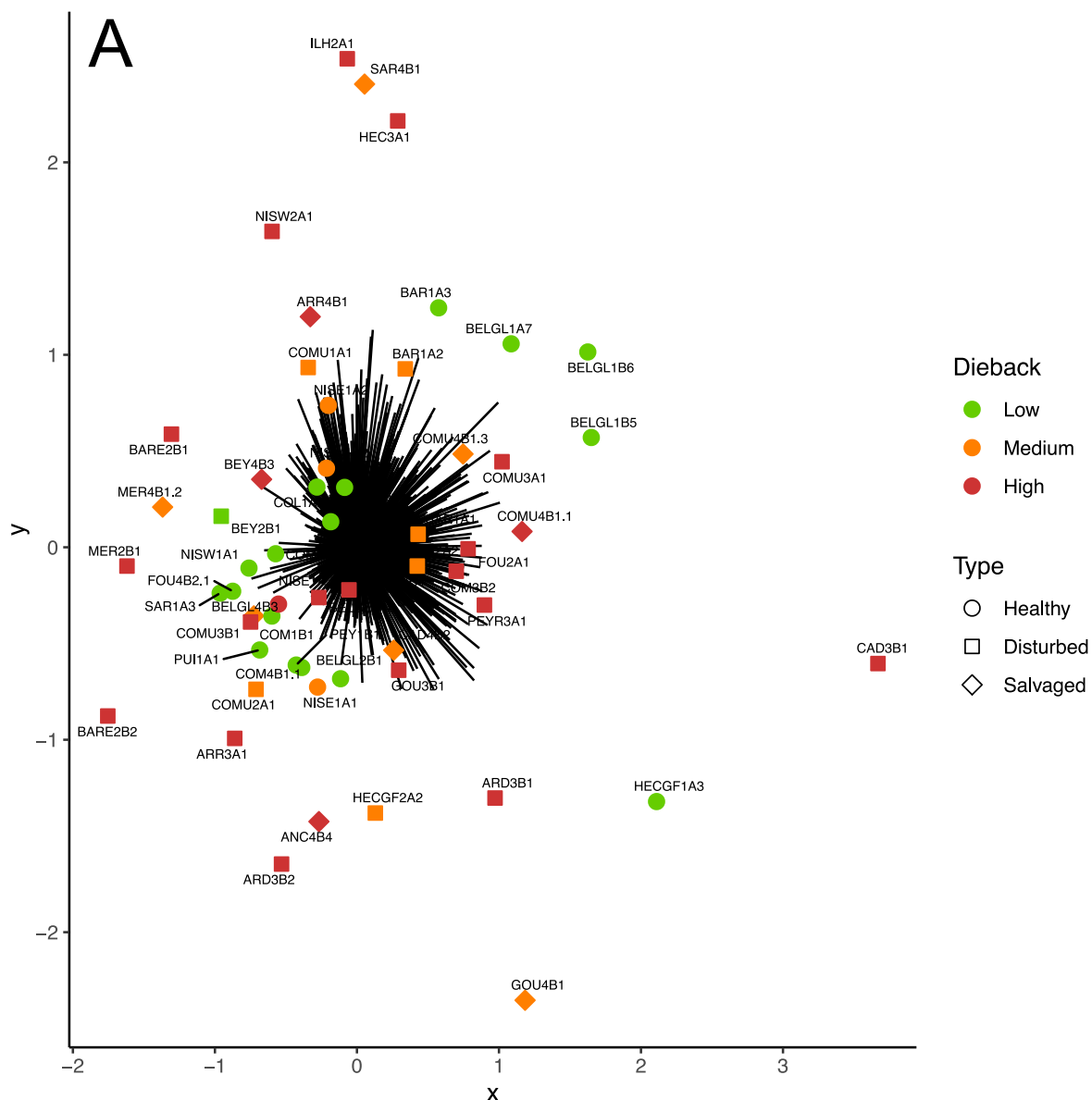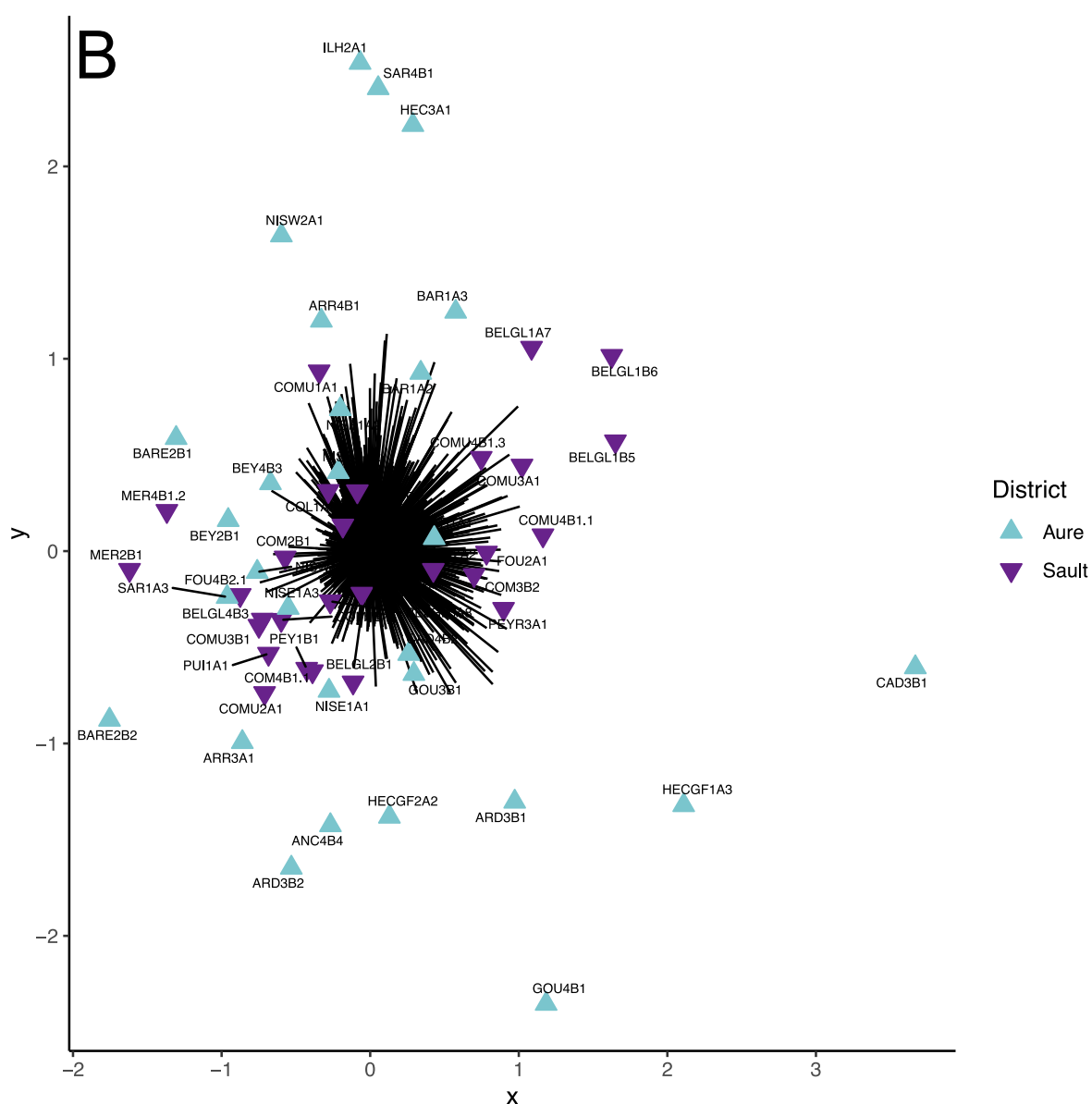
Fig. 5: **Effect of dieback on community composition for insect orders and functions.** Representation of the species retention rate (*i.e.* zeta ratio) per plot (*i.e.* zeta order) following nearest-neighbour plot combinations scheme (NN) for low, medium and high dieback levels, respectively (**A, B, C**) for the five main insect Orders (Coleoptera, Diptera, Hemiptera, Hymenoptera and Lepidoptera) and (**D, E, F**) for the four main ecological functions recovered from taxonomic assignment (floricolous / non-floricolous and parasitoid / non-parasitoid species). Green line with plain dots represents mean species retention rate of the total dataset in each respective dieback category. Increasing curves express that common MOTUs are more likely to be retained in additional samples than rare ones (with presence of common species over all plots if zeta ratio = 1) and decreasing curves indicates species turnover. In our study, no core of common species could be sampled for Coleoptera at low dieback level as well as for Lepidoptera and Hemiptera throughout the entire dataset for each dieback category, respectively. For both Diptera and Hymenoptera, common core of species was observed within all dieback level plots except between high dieback level ones. Similarly, drop and lack of structure in common species was detected for floricolous and parasitoid functional assemblages within high dieback level plots while stable for the other two dieback levels, with this functional turnover being driven by dipteran and hymenopteran species turnovers at high dieback.

846

Fig. 6: **Effect of stand type on community composition for insect orders and functions.**
Representation of the species retention rate (*i.e.* zeta ratio) per plot (*i.e.* zeta order) following nearest-neighbour plot combinations scheme (NN) for healthy, disturbed and salvaged stands, respectively (**A, B, C**) for the five main insect Orders (Coleoptera, Diptera, Hemiptera, Hymenoptera and Lepidoptera) and (**D, E, F**) for the four main ecological functions recovered from taxonomic assignment (floricolous / non-floricolous and parasitoid / non-parasitoid species). Green line with plain dots represents mean species retention rate of the total dataset in each respective dieback category. Increasing curves express that common MOTUs are more likely to be retained in additional samples than rare ones (with presence of common species over all plots if zeta ratio = 1) and decreasing curves indicates species turnover. In the present case, we found no core of common species for Hemiptera within each studied stand type. For clarity purposes, Hemiptera are hereafter not considered. Within healthy stands, only Coleoptera had no common species retained, while Diptera had a slight drop in the probability of common species occurring, but yet not a complete compositional turnover. Lepidoptera, Diptera and Hymenoptera were all impacted by disturbed stands with complete compositional turnovers within studied plots, and similar observation could be noticed for parasitoid and floricolous functional assemblages. Meanwhile Coleoptera had a constant core of common species across the considered plots, and similar observation could be made within salvaged plots. Again, Lepidoptera assemblages were fully heterogenous across salvaged logged plots, and both Diptera and Hymenoptera expressed a lower probability of sampling common species, but yet not showing a complete change in community composition.

32

869
870
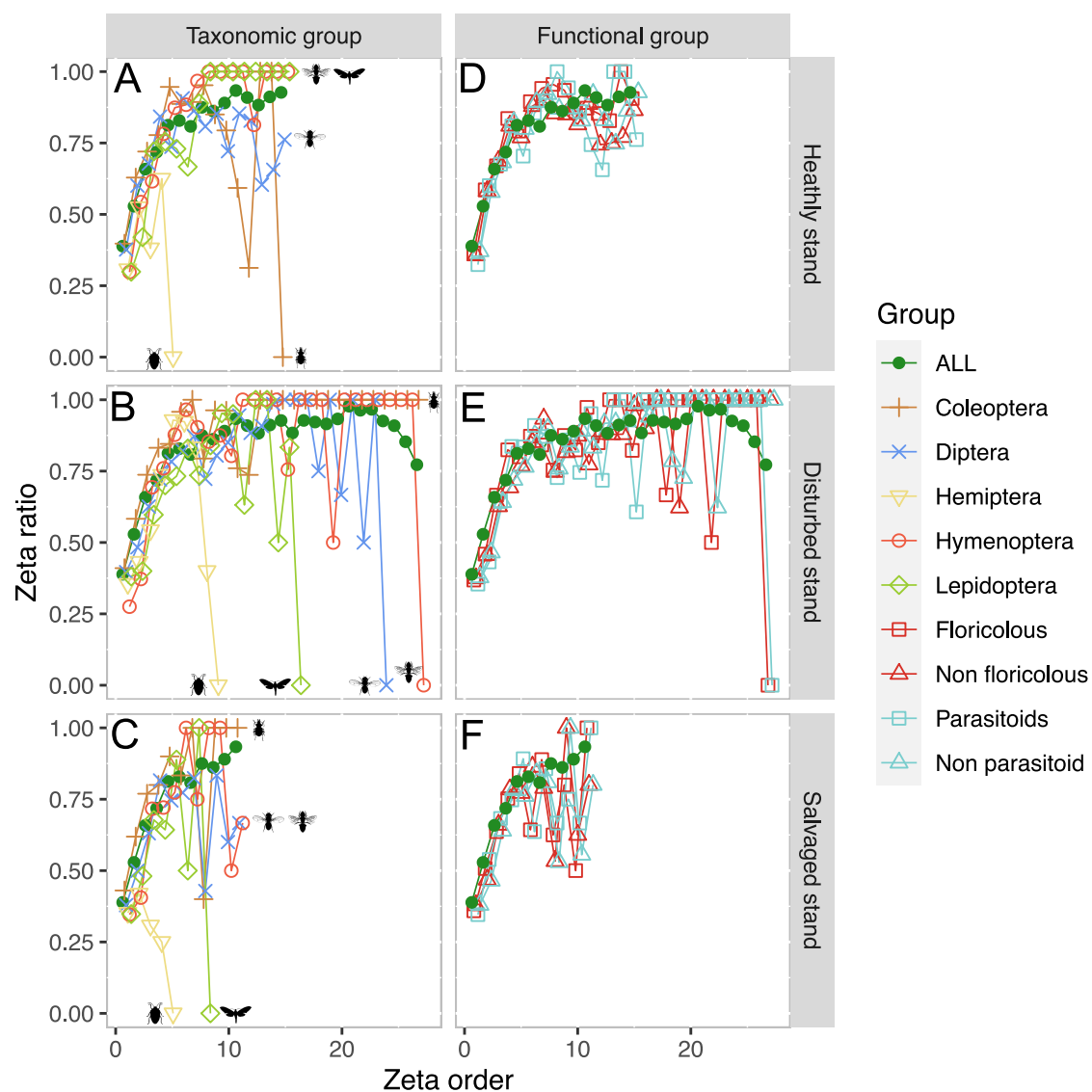
871 <u>Fig. 7</u>: **Insect orders winning and losing from environmental conditions.**
872 Heatmap representation based from traitGLM analyses of insect orders' prevalence according
873 to environmental conditions: (**A**) forest dieback level (*i.e.* low, medium and high) and (**B**) stand
874 types (*i.e.* healthy, disturbed and salvaged). Negative and positive interactions and their
875 intensity in terms of biodiversity response to environmental conditions for each of the five most
876 represented orders are highlighted by a continuous colour gradient spanning from blue to red,
877 respectively. We found specific responses for each insect orders' diversity in regards to dieback
878 levels and stand types. For instance, the low prevalence of dipterans and hymenopterans
879 observed at low dieback level while relatively high prevalence for coleopterans. However,
880 hymenopteran prevalence was greater in forests of medium dieback level. Coleoptera was the
881 only order with prevalence negatively impacted by salvage logging. Finally, disturbed plots
882 showed a relatively high prevalence of both Lepidoptera and Hemiptera.

883

884
885
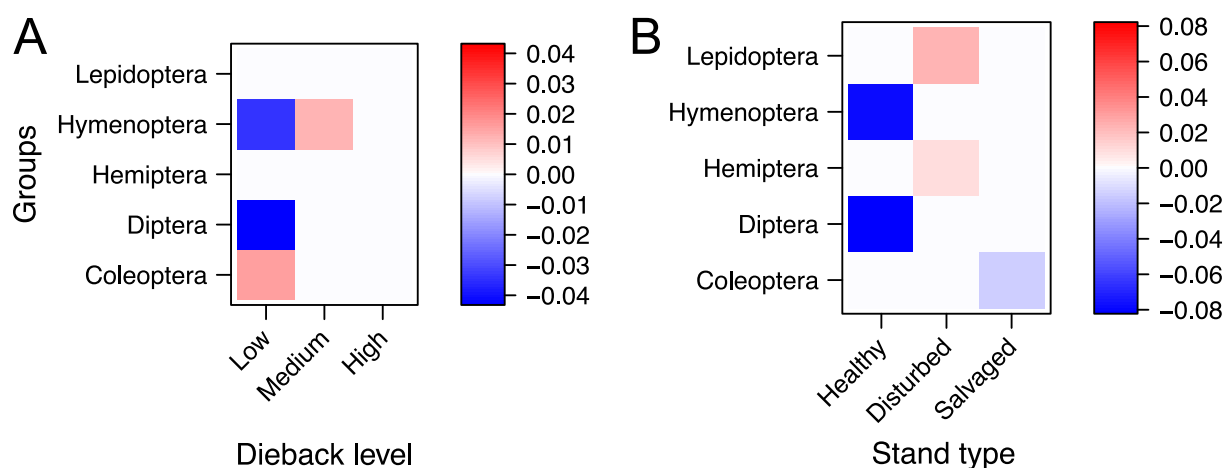
Fig. 8: **Environmental variables driving species community changes (zeta diversity changes).**

(**A**) I-splines values of predictors for seven environmental variables with their respective contributions to zeta diversity. Environmental contributions are shown for six different zeta orders (*i.e.* ζ_i) here referring to plots compared together (from $\zeta2$ to $\zeta50$), with the main contributively environmental variables highlighted for each zeta order. We observed a shift in the key forest features explaining community changes according to geographic scale, with canopy openness and altitude as main environmental drivers at local scale ($\zeta2$), replaced by tree-related microhabitats and deadwood amount at large scales (from $\zeta28$ to $\zeta50$). (**B**) Proportion (in r^2) of the total zeta diversity variance explained by environmental variables, distance, undistinguishable variables or unexplained, according to zeta order. The seven tested environmental variables accounted for 15–20% or the total variance in community changes regardless of zeta order, with distance explaining up to 20% of the total variance as well depending on the different zeta orders considered.
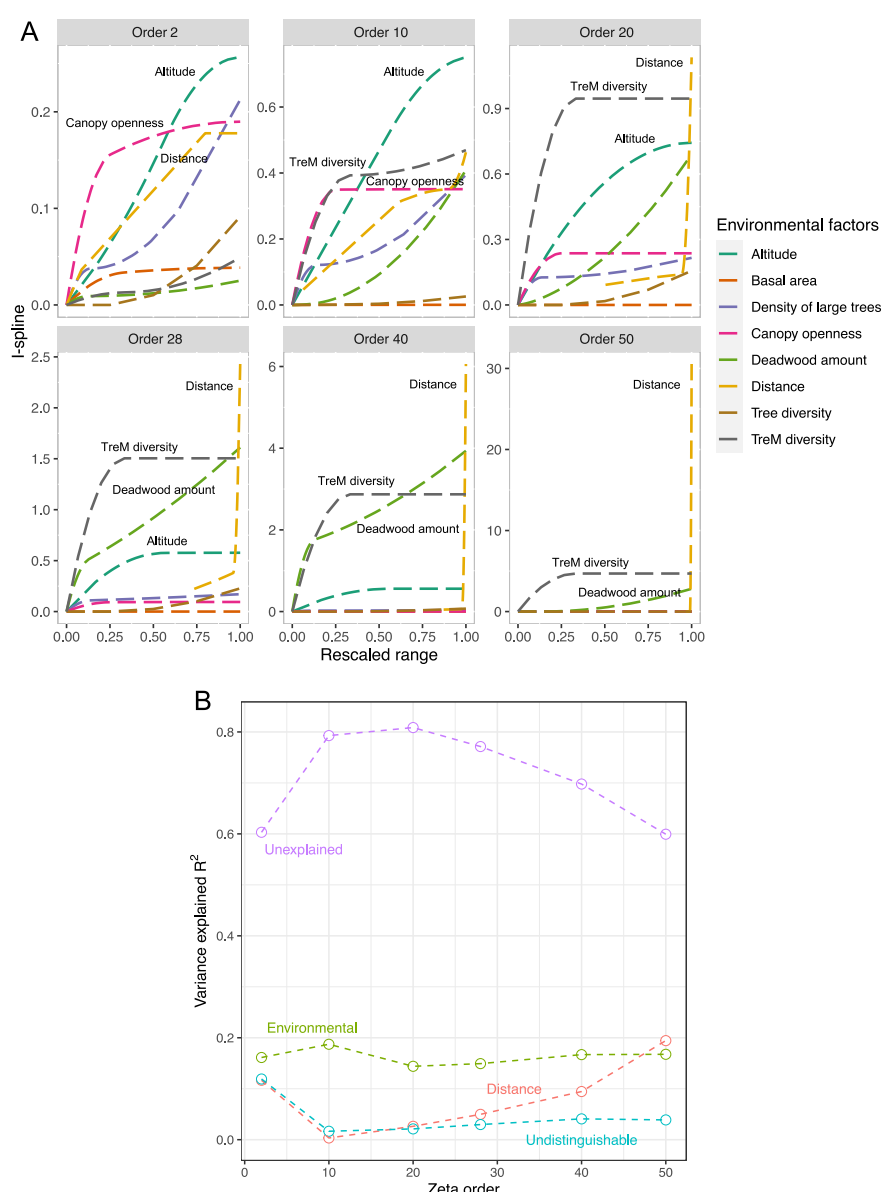
Table I: **Impact of forest dieback and salvage logging on insect species richness.**
Generalized linear models for species richness variations of different study groups (*i.e.* total insects, five individual insect orders, and four functional group) compared across respective environmental conditions of diebacks (*i.e.* low, medium and high forest dieback levels) and stand types (*i.e.* healthy, disturbed and salvaged logged). Functional groups (*i.e.* Parasitoids/non-parasitoids, floricolous/non-floricolous insects) were assigned using each MOTU's taxonomic family. Significance is given by "*" while "*N.S.*" stands for non-significant.

Studied group	Condition	Degree of freedom	Chi-square	*P*-value (95% confidence)	Significance
Total insects					
	Dieback level effect	2	2.10963	0.3483	*N.S.*
	Stand type effect	2	0.94467	0.623	*N.S.*
	District effect	1	0.51971	0.4710	*N.S.*
Coleoptera					
	Dieback level effect	2	0.3568	0.8365938	*N.S.*
	Stand type effect	2	0.9060	0.6357279	*N.S.*
	District effect	1	13.9565	**0.0001871**	***
Diptera					
	Dieback level effect	2	1.9782	0.3719	*N.S.*
	Stand type effect	2	1.8063	0.4053	*N.S.*
	District effect	1	1.8153	0.1779	*N.S.*
Hemiptera					
	Dieback level effect	2	0.14194	0.9315	*N.S.*
	Stand type effect	2	1.64021	0.4404	*N.S.*
	District effect	1	1.58658	0.2078	*N.S.*
Hymenoptera					
	Dieback level effect	2	2.7238	0.2562	*N.S.*
	Stand type effect	2	1.2385	0.5384	*N.S.*
	District effect	1	1.0418	0.3074	*N.S.*

Lepidoptera

		df			
	Dieback level effect	*2*	0.74240	0,6899	*N.S.*
	Stand type effect	*2*	0.14966	0.9279	*N.S.*
	District effect	1	0.92520	0.3361	*N.S.*

Floricolous

	Dieback level effect	*2*	2.33042	0.3119	*N.S.*
	Stand type effect	*2*	0.37240	0.8301	*N.S.*
	District effect	1	0.57194	0.4495	*N.S.*

Non-floricolous

	Dieback level effect	*2*	2.10039	0.3499	*N.S.*
	Stand type effect	*2*	1.34128	0.5114	*N.S.*
	District effect	1	0.10885	0.7415	*N.S.*

Parasitoids

	Dieback level effect	*2*	2.4223	0.29785	*N.S.*
	Stand type effect	*2*	0.7983	0.67089	*N.S.*
	District effect	1	3.2229	0.007261	*N.S.*

Non-parasitoids

	Dieback level effect	*2*	2.01770	0.3646	*N.S.*
	Stand type effect	*2*	1.13627	0.5666	*N.S.*
	District effect	1	0.00031	0.9860	*N.S.*

908

References

1. Stork, N. E. How Many Species of Insects and Other Terrestrial Arthropods Are There on Earth? *Annu. Rev. Entomol.* **63**, 31–45 (2018). https://doi.org/10.1146/annurev-ento-020117-043348

2. Seibold, S. *et al.* Arthropod decline in grasslands and forests is associated with landscape-level drivers. *Nature* **574**, 671–674 (2019). https://doi.org/10.1038/s41586-019-1684-3

3. Wagner, D. L., Fox, R., Salcido, D. M. & Dyer, L. A. A window to the world of global insect declines: Moth biodiversity trends are complex and heterogeneous. *Proc. Natl. Acad. Sci.* **118** (2), e2002549117, 1–8 (2021). https://doi.org/10.1073/pnas.2002549117

4. Kulakowski, D. *et al.* A walk on the wild side: Disturbance dynamics and the conservation and management of European mountain forest ecosystems. *For. Ecol. Manag.* **388**, 120–131 (2017). https://doi.org/10.1016/j.foreco.2016.07.037

5. Allen, C. D., Breshears, D. D. & McDowell, N. G. On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere* **6** (8), 129, 1–55 (2015). https://doi.org/10.1890/ES15-00203.1

6. Senf, C., Buras, A., Zang, C. S., Rammig, A. & Seidl, R. Excess forest mortality is consistently linked to drought across Europe. *Nat. Commun.* **11**, 6200, 1–8 (2020). https://doi.org/10.1038/s41467-020-19924-1

7. Brodribb, T. J., Powers, J., Cochard, H. & Choat, B. Hanging by a thread? Forests and drought. *Science* **368**, (6488), 261–266 (2020). DOI: 10.1126/science.aat7631

8. Samaniego, L. *et al.* Anthropogenic warming exacerbates European soil moisture droughts. *Nat. Clim. Change* **8**, 421–426 (2018). https://doi.org/10.1038/s41558-018-0138-5

9. Sallé, A. *et al.* Forest decline differentially affects trophic guilds of canopy-dwelling beetles. *Annals of Forest Science* **77**, 86 (2020) https://doi.org/10.1007/s13595-020-00990-w

10. Beudert, B. *et al.* Bark Beetles Increase Biodiversity While Maintaining Drinking Water Quality: Bark beetles, biodiversity and drinking water. *Conserv. Lett.* **8** (4), 272–281 (2015). https://doi.org/10.1111/conl.12153

11. Evans, P. M. *et al.* Testing the relative sensitivity of 102 ecological variables as indicators of woodland condition in the New Forest, UK. *Ecol. Indic.* **107**, 105575, 1–12 (2019). https://doi.org/10.1016/j.ecolind.2019.105575

12. Seidl, R. *et al.* Forest disturbances under climate change. *Nat. Clim. Change* **7**, 395–402 (2017). https://doi.org/10.1038/nclimate3303

13. Thorn, S. *et al.* Impacts of salvage logging on biodiversity: A meta-analysis. *J. Appl. Ecol.* **55**, 279–289 (2018). https://doi.org/10.1111/1365-2664.12945

14. Thom, D. & Seidl, R. Natural disturbance impacts on ecosystem services and biodiversity in temperate and boreal forests. *Biol. Rev.* **91**, 760–781 (2016). https://doi.org/10.1111/brv.12193

15. Moretti, M., Obrist, M. K. & Duelli, P. Arthropod biodiversity after forest fires: winners and losers in the winter fire regime of the southern Alps. *Ecography* **27** (2), 173–186 (2004). https://doi.org/10.1111/j.0906-7590.2004.03660.x

16. Dainese, M. *et al.* A global synthesis reveals biodiversity-mediated benefits for crop production. *Sci. Adv.* **5** (10), eaax0121, 1–13 (2019). DOI: 10.1126/sciadv.aax0121

17. Larsen, S., Chase, J. M., Durance, I. & Ormerod, S. J. Lifting the veil: richness measurements fail to detect systematic biodiversity change over three decades. *Ecology* **99** (6), 1316–1326 (2018). https://doi.org/10.1002/ecy.2213

18. Cardoso, P. The seven impediments in invertebrate conservation and how to overcome them. *Biol. Conserv.* **144** (11), 2647–2655 (2011). https://doi.org/10.1016/j.biocon.2011.07.024

19. Morinière, J. *et al.* A DNA barcode library for 5,200 German flies and midges (Insecta: Diptera) and its implications for metabarcoding-based biomonitoring. *Mol. Ecol. Resour.* **19** (4), 900–928 (2019). https://doi.org/10.1111/1755-0998.13022

20. Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. Biological Identifications through DNA Barcodes. *Proc. Biol. Sci.* **270** (1512), 313–321 (2003). https://doi.org/10.1098/rspb.2002.2218

21. Ratnasingham, S. & Hebert, P. D. N. A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. PLoS ONE **8** (7), e66213, 1–16 (2013). https://doi.org/10.1371/journal.pone.0066213

22. Wang, C. *et al.* The ecological impact of pest-induced tree dieback on insect biodiversity in Yunnan pine plantations, China. *For. Ecol. Manag.* **491**, 119173, 1–11 (2021). https://doi.org/10.1016/j.foreco.2021.119173

23. Péré, C., Jactel, H. & Kenis, M. Response of insect parasitism to elevation depends on host and parasitoid life-history strategies. *Biol. Lett.* **9** (4), 20130028, 1–4 (2013). https://doi.org/10.1098/rsbl.2013.0028

24. Nguyen, L.-T.-H. *et al.* Bottom-up effect of water stress on the aphid parasitoid Aphidius ervi. *Entomol. Gen.* **38**, 15–27 (2018). DOI: 10.1127/entomologia/2018/0575

25. Lebourgeois, F., Rathgeber, C. B. K. & Ulrich, E. Sensitivity of French temperate coniferous forests to climate variability and extreme events (*Abies alba*, *Picea abies* and *Pinus sylvestris*). *J. Veg. Sci.* **21** (2), 364–376 (2010). https://doi.org/10.1111/j.1654-1103.2009.01148.x

26. Camarero, J. J., Bigler, C., Linares, J. C. & Gil-Pelegrín, E. Synergistic effects of past historical logging and drought on the decline of Pyrenean silver fir forests. *For. Ecol. Manag.* **262** (5), 759–769 (2011). https://doi.org/10.1016/j.foreco.2011.05.009

27. Ratnasingham, S. & Hebert, P. D. N. BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). *Mol. Ecol. Notes* **7** (3), 355–364 (2007). https://doi.org/10.1111/j.1471-8286.2007.01678.x

28. Hui, C. & McGeoch, M. A. Zeta Diversity as a Concept and Metric That Unifies Incidence-Based Biodiversity Patterns. *Am. Nat.* **184** (5), 684–694 (2014). https://doi.org/10.1086/678125

29. McGeoch, M. A. *et al.* Measuring continuous compositional change using decline and decay in zeta diversity. *Ecology* **100** (11), e02832, 1–18 (2019). https://doi.org/10.1002/ecy.2832

30. Latombe, G., McGeoch, M., Nipperess, D. & Hui, C. zetadiv: Functions to Compute Compositional Turnover Using Zeta Diversity. (version 1.2.0) [software] available at https://github.com/GLatombe/zetadiv (2020).

31. Brown, A. M. *et al.* The fourth-corner solution - using predictive models to understand how species traits interact with the environment. *Methods Ecol. Evol.* **5** (4), 344–352 (2014). https://doi.org/10.1111/2041-210X.12163

32. Hartop, E., Srivathsan, A., Ronquist, F. & Meier, R. Large-scale Integrative Taxonomy (LIT): resolving the data conundrum for dark taxa. Preprint at *BioRxiv* (2016). https://doi.org/10.1101/2021.04.13.439467

33. Kortmann, M. *et al.* Ecology versus society: Impacts of bark beetle infestations on biodiversity and restorativeness in protected areas of Central Europe. *Biol. Conserv.* **254**, 10893, 1–9 (2021). https://doi.org/10.1016/j.biocon.2020.108931

34. Thorn, S. *et al.* Rare species, functional groups, and evolutionary lineages drive successional trajectories in disturbed forests. *Ecology* **101** (3), e02949, 1–8 (2020). https://doi.org/10.1002/ecy.2949

35. Müller, J., Noss, R. F., Bussler, H. & Brandl, R. Learning from a "benign neglect strategy" in a national park: Response of saproxylic beetles to dead wood accumulation. *Biol. Conserv.* **143** (11), 2559–2569 (2010). https://doi.org/10.1016/j.biocon.2010.06.024

36. Cours, J. *et al.* Contrasting responses of habitat conditions and insect biodiversity to pest- or climate-induced dieback in coniferous mountain forests. *For. Ecol. Manag.* **482**, 11881, 1–14 (2021). https://doi.org/10.1016/j.foreco.2020.118811

37. Herrault, P.-A. *et al.* Combined effects of area, connectivity, history and structural heterogeneity of woodlands on the species richness of hoverflies (Diptera: Syrphidae). *Landsc. Ecol.* **31**, 877–893 (2016). DOI 10.1007/s10980-015-0304-3

38. Leather, S. R. "Ecological Armageddon" - more evidence for the drastic decline in insect numbers: Insect declines. *Ann. Appl. Biol.* **172** (1), 1–3 (2018). https://doi.org/10.1111/aab.12410

39. Fleishman, E., Noss, R. & Noon, B. Utility and limitations of species richness metrics for conservation planning. *Ecol. Indic.* **6** (3), 543–553 (2006). https://doi.org/10.1016/j.ecolind.2005.07.005

40. Borges, P. A. V., Rigal, F., Ros-Prieto, A. & Cardoso, P. Increase of insular exotic arthropod diversity is a fundamental dimension of the current biodiversity crisis. *Insect Conserv. Divers.* **13** (5), 508–518 (2020). https://doi.org/10.1111/icad.12431

41. Ienco, A., Dapporto, L., Greco, S., Infusino, M. & Scalercio, S. β-Diversity partitioning of moth communities within and between different forest types. *Sci. Nat.* **107**, 8, 1–11 (2020). https://doi.org/10.1007/s00114-020-1665-2

42. Thorn, S. *et al.* The living dead: acknowledging life after tree death to stop forest degradation. *Front. Ecol. Environ.* **18** (9), 505–512 (2020) https://doi.org/10.1002/fee.2252

43. Arnan, X., Bosch, J., Comas, L., Gracia, M. & Retana, J. Habitat determinants of abundance, structure and composition of flying Hymenoptera communities in mountain old-growth forests: Habitat determinants of flying Hymenoptera. *Insect Conserv. Divers.* **4** (3), 200–211 (2011). https://doi.org/10.1111/j.1752-4598.2010.00123.x

44. Säterberg, T., Jonsson, T., Yearsley, J., Berg, S. & Ebenman, B. A potential role for rare species in ecosystem dynamics. *Sci. Rep.* **9**, 11107, 1–12 (2019). https://doi.org/10.1038/s41598-019-47541-6

45. Jain, M. *et al.* The importance of rare species: a trait-based assessment of rare species contributions to functional diversity and possible ecosystem function in tall-grass prairies. *Ecol. Evol.* **4**, 104–112 (2014). https://doi.org/10.1002/ece3.915

46. Soliveres, S. *et al.* Locally rare species influence grassland ecosystem multifunctionality. *Philos. Trans. R. Soc. B Biol. Sci.* **371** (1694), 20150269, 1–10 (2016). https://doi.org/10.1098/rstb.2015.0269

47. Heidrich, L. *et al.* Heterogeneity–diversity relationships differ between and within trophic levels in temperate forests. *Nat. Ecol. Evol.* **4**, 1204–1212 (2020). https://doi.org/10.1038/s41559-020-1245-z

48. Larrieu, L. *et al*. Tree related microhabitats in temperate and Mediterranean European forests: A hierarchical typology for inventory standardization. *Ecol. Indic*. **84**, 194–207 (2018). https://doi.org/10.1016/j.ecolind.2017.08.051

49. Seibold, S. *et al*. Microclimate and habitat heterogeneity as the major drivers of beetle diversity in dead wood. *J. Appl. Ecol*. **53** (3), 934–943 (2016). https://doi.org/10.1111/1365-2664.12607

50. Sire, L. *et al.* The Challenge of DNA Barcoding Saproxylic Beetles in Natural History Collections—Exploring the Potential of Parallel Multiplex Sequencing With Illumina MiSeq. *Front. Ecol. Evol.* **7**, 495, 1–12 (2019). https://doi.org/10.3389/fevo.2019.00495

51. Morinière, J. *et al.* Species Identification in Malaise Trap Samples by DNA Barcoding Based on NGS Technologies and a Scoring Matrix. *PLoS ONE* **11** (5), e0155497, 1–14 (2016). https://doi.org/10.1371/journal.pone.0155497

52. Ashfaq, M. *et al.* Insect diversity in the Saharo-Arabian region: Revealing a little-studied fauna by DNA barcoding. *PLoS ONE* **13** (7), e0199965, 1–16 (2018). https://doi.org/10.1371/journal.pone.0199965

53. Karlsson, D., Hartop, E., Forshage, M., Jaschhof, M. & Ronquist, F. The Swedish Malaise Trap Project: A 15 Year Retrospective on a Countrywide Insect Inventory. *Biodivers. Data J.* **8**, e47255, 1–35 (2020). doi: 10.3897/BDJ.8.e47255

54. Russo, L., Stehouwer, R., Heberling, J. M. & Shea, K. The Composite Insect Trap: An Innovative Combination Trap for Biologically Diverse Sampling. *PLoS ONE* **6** (6), e21079, 1–7 (2011). https://doi.org/10.1371/journal.pone.0021079

55. Lambert, J., Drenou, C., Denux, J.-P., Balent, G. & Cheret, V. Monitoring forest decline through remote sensing time series analysis. *GIScience Remote Sens*. **50**, 437–457 (2013). https://doi.org/10.1080/15481603.2013.820070

56. Lemmon, P. E. A spherical densiometer for estimating forest overstory density. *For. Sci.* **2** (4), 314–320 (1956). https://doi.org/10.1093/forestscience/2.4.314

57. Larrieu, L. & Gonin, P. L'indice de biodiversité potentielle (ibp) : une méthode simple et rapide pour évaluer la biodiversité potentielle des peuplements forestiers. *Rev. For. Fr.* **6**, 727–748 (2008). https://doi.org/10.4267/2042/28373

58. Larsson, T.-B. Biodiversity Evaluation Tools for European Forests. 75–81, in *Criteria and Indicators for Sustainable Forest Management at the Forest Management Unit Level*, 280pp. Eds. Franc, A., Laroussinie, O., Karjalainen, T. (European Forest Institute proceeding vol. 38, Nancy, 2001). ISBN: 952-9844-86-7

59. Gosselin, F. & Larrieu, L. Developing and using statistical tools to estimate observer effect for ordered class data: The case of the IBP (Index of Biodiversity Potential). *Ecol. Indic.* **110**, 105884, 1–10 (2020). https://doi.org/10.1016/j.ecolind.2019.105884

60. Elbrecht, V., Peinert, B. & Leese, F. Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecol. Evol*. **7** (17), 6918–6926 (2017). https://doi.org/10.1002/ece3.3192

61. Leray, M., Haenel, Q. & Bourlat, S. J. Preparation of Amplicon Libraries for Metabarcoding of Marine Eukaryotes Using Illumina MiSeq: The Adapter Ligation Method. 209–218 in *Marine Genomics* Ed. Bourlat, S. J. (Methods in Molecular Biology vol. 1452, Springer New York, New York, 2016). https://doi.org/10.1007/978-1-4939-3774-5_14

62. Leray, M. *et al*. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front. Zool.* **10**, 34, 1–14 (2013). https://doi.org/10.1186/1742-9994-10-34

63. Geller, J., Meyer, C., Parker, M. & Hawk, H. Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Mol. Ecol. Resour.* **13** (5), 851–861 (2013). https://doi.org/10.1111/1755-0998.12138

64. Fadrosh, D. W. *et al.* An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* **2**, 6, 1–7 (2014). https://doi.org/10.1186/2049-2618-2-6

65. Illumina. Illumina Two-Channel SBS Sequencing Technology. Pub. No. 770-2013-054. (2016). Available at https://www.well.ox.ac.uk/ogc/wp-content/uploads/2017/09/techspotlight_two-channel_sbs.pdf

66. Knittel, T. & Picard, D. PCR with degenerate primers 9 containing deoxyinosine fails with PFU DNA polumerase. *Genome Res.* **2**, 346–347 (1993). doi:10.1101/gr.2.4.346

67. Alberdi, A., Aizpurua, O., Gilbert, M. T. P. & Bohmann, K. Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods Ecol. Evol.* **9**, 134–147 (2018). https://doi.org/10.1111/2041-210X.12849

68. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421, 1–10 (2009). https://doi.org/10.1186/1471-2105-10-421

69. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88, 1–7 (2016). https://doi.org/10.1186/s13104-016-1900-2

70. Joshi, N. A. & Fass, J. N. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. (version 1.33) [software] available at https://github.com/najoshi/sickle (2011).

71. Nurk, S. *et al.* Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. 158–170, in *RECOMB 2013: Research in Computational Molecular Biology*, 346pp. Eds. Deng, M., Jiang, R., Sun, F. & Zhang, X. part of *Lecture Notes in Computer Science* (Springer International Publishing, vol. 7821, Berlin, Heidelberg, 2013). https://doi.org/10.1007/978-3-642-37195-0_13

72. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* **13**, 31, 1–7 (2012). https://doi.org/10.1186/1471-2105-13-31

73. Zepeda-Mendoza, M. L., Bohmann, K., Carmona Baez, A. & Gilbert, M. T. P. DAMe: a toolkit for the initial processing of datasets with PCR replicates of double-tagged amplicons for DNA metabarcoding analyses. *BMC Res. Notes* **9**, 255, 1–13 (2016). https://doi.org/10.1186/s13104-016-2064-9

74. R Core Team. R: A language and environment for statistical computing. (R Foundation for Statistical Computing, Vienna, Austria, 2017). https://www.R-project.org/

75. Renkonen, O. Statistisch-ökologische Untersuchungen über die terrestische Käferwelt der finnischen Bruchmoore. *Ann Bot Soc Zool-Bot Fenn Vanamo* **6**, 1–231 (1938).

76. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584, 1–22 (2016). DOI: 10.7717/peerj.2584

77. Hebert, P. D. N., Ratnasingham, S. & de Waard, J. R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. B Biol. Sci.* **270**, S96–S99 (2003). https://doi.org/10.1098/rsbl.2003.0025

78. Boyer, F. *et al.* OBITOOLS: a UNIX-inspired software package for DNA metabarcoding. *Mol. Ecol. Resour.* **16**, 176–182 (2016). https://doi.org/10.1111/1755-0998.12428

79. Frøslev, T. G. *et al.* Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nat. Commun.* **8**, 1188, 1–11 (2017). https://doi.org/10.1038/s41467-017-01312-x

80. Chamberlain, S. bold: Interface to Bold Systems API. (version 1.1.0) [software] (2020). https://CRAN.R-project.org/package=bold

81. Godfray, C. H. J. Parasitoids: Behavioral and Evolutionary Ecology. 472 pp. (Princeton University Press, Princeton, New Jersey, 1994). ISBN: 0-691-00047-6

82. Bouget, C., Brustel, H. & Zagatti, P. The French Information System on Saproxylic BEetle Ecology (FRISBEE): an ecological and taxonomical database to help with the assessment of forest conservation status. *Rev. D'Ecologie Terre Vie Société Natl. Prot. Nat*. 33–36 (2008). (hal-00454436)

83. Schmidt-Kloiber, A. & Hering, D. www.freshwaterecology.info – An online tool that unifies, standardises and codifies more than 20,000 European freshwater organisms and their ecological preferences. *Ecol. Indic.* **53**, 271–282 (2015). https://doi.org/10.1016/j.ecolind.2015.02.007

84. Liu, M., Clarke, L. J., Baker, S. C., Jordan, G. J. & Burridge, C. P. A practical guide to DNA metabarcoding for entomological ecologists. *Ecol. Entomol.* **45** (3), 373–385 (2020). https://doi.org/10.1111/een.12831

85. Hsieh, T. C., Ma, K. H. & Chao, iNEXT: iNterpolation and EXTrapolation for species diversity. (version 2.0.20) [software] (2020). http://chao.stat.nthu.edu.tw/wordpress/software-download/

86. Oksanen, J. *et al.* vegan: Community Ecology Package. (version 2.5-6) [software] (2020). https://CRAN.R-project.org/package=vegan

87. Fox, J. & Weisberg, S. An R Companion to Applied Regression. (SAGE Publications, London, United Kingdom, 2019). ISBN: 978-1-5443-3647-3

88. Wang, Y., Naumann, U., Eddelbuettel, D., Wilshire, J. & Warton, D. mvabund: Statistical Methods for Analysing Multivariate Abundance Data. (version 4.3.1) [software] (2020). https://CRAN.R-project.org/package=mvabund

89. Popovic, G. C., Warton, D. I., Thomson, F. J., Hui, F. K. C. & Moles, A. T. Untangling direct species associations from indirect mediator species effects with graphical models. *Methods Ecol. Evol.* **10** (9), 1571–1583 (2019). https://doi.org/10.1111/2041-210X.13247

90. De Cáceres, M., Legendre, P. & Moretti, M. Improving indicator species analysis by combining groups of sites. *Oikos* **119** (10), 1674–1684 (2010). https://doi.org/10.1111/j.1600-0706.2010.18334.x

91. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8** (4), e61217, 1–11 (2013). https://doi.org/10.1371/journal.pone.0061217

92. Gaston, K. J. What is Rarity? 1–21, in *Rarity*, 205pp, part of *Population and Community Biology Series* (Springer, vol. 13, Dordrecht, 1994). https://doi.org/10.1007/978-94-011-0701-3_1

93. Akaike, H. Information theory and an extension of the maximum likelihood principle. in *Second International Symposium on Information Theory*, 267–281 (1973).

94. Ferrier, S., Manion, G., Elith, J. & Richardson, K. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Divers. Distrib*. **13** (3), 252–264 (2007). https://doi.org/10.1111/j.1472-4642.2007.00341.x