Cryptic diversity and population structure at small scales: The freshwater snail *Ancylus* (Planorbidae, Pulmonata) in the Montseny mountain range

Author names and affiliations

Jan N. Macher[1*], Martina Weiss[1], Arne J. Beermann[1] & Florian Leese[1]

[1] Aquatic Ecosystem Research, University of Duisburg-Essen, Universitätsstraße 5, 45141 Essen, Germany

***Corresponding author:** jan.macher@uni-due.de

E-mail: Martina Weiss: Martina.Weiss@uni-due.de

Arne Beermann: Arne.Beermann@uni-due.de

Florian Leese: Florian.Leese@uni-due.de

ABSTRACT

Anthropogenic impacts like intensified land use and climate change are severe threats to freshwater biodiversity and effective biodiversity monitoring is therefore one of the most urgent tasks. This is however often hampered by the lack of knowledge regarding the number and ecology of species. Molecular tools have shown many freshwater taxa to comprise morphologically cryptic species, which often occur in sympatry on a small geographic scale. Here, we studied the freshwater snail *Ancylus fluviatilis* (MÜLLER, 1774) species complex in the Iberian Montseny Mountains. We hypothesised 1) that several species of *A. fluviatilis* sensu lato occur in the Montseny, 2) that different *Ancylus* species seldom co-occur in syntopy due to different ecological demands or interspecific competition, and 3) that species show a pattern of strong population structure within streams or catchments due to ecological preferences or local adaptation. We barcoded 180 specimens from 36 sites in the Montseny for the cytochrome c oxidase subunit I (COI) barcoding gene and molecularly identified two *Ancylus* species. These species seldom occurred in syntopy and a species distribution modelling approach showed differing bioclimatic preferences of the species. One species mainly occurs in cooler, higher altitude streams while the second species occurs in lower-altitude areas with higher temperatures. Tests of population structure showed that both species possibly do not disperse well in the study area and that populations within species are likely adapted to certain bioclimatic conditions in different regions of the Montseny. Our results highlight the need to incorporate molecular techniques into routine monitoring programmes.

Keywords: cryptic species complex, population structure, barcoding, species distribution modelling, freshwater invertebrates

INTRODUCTION

Anthropogenic impacts like intensified land use and climate change are severe threats to biodiversity (Vörösmarty *et al.*, 2010; Steffen *et al.*, 2015). Therefore, monitoring the ongoing loss of biodiversity is highly important and many countries worldwide have established programmes to do so. However, the total loss of biodiversity can only be monitored when accurate knowledge on the number, ecology, distribution and genetic diversity of species is available. For many areas and ecosystems, such information does often either not exist or is inaccurate: in recent years, the use of molecular methods has shown that the number of species is underestimated in many taxa (e.g. Amato *et al.*, 2007; Pfenninger and Schwenk, 2007; Adams *et al.*, 2014). This is especially true for freshwater ecosystems, which harbour a large number of morphologically indistinguishable or cryptic animal species (e.g. Pauls *et al.*, 2010; Weigand *et al.*, 2011; Weiss *et al.*, 2014). The ecology of most cryptic species is however rarely known since it has been studied for relatively few taxa only (e.g. Ortells *et al.*, 2003; Rissler and Apodaca, 2007; Lagrue *et al.*, 2014; Fišer *et al.*, 2015). This lack of knowledge poses a risk, since monitoring programmes and biodiversity assessments can come to wrong conclusions if species with different ecologies are treated as being identical regarding their ecological demands and thus, their suitability to indicate ecosystem health (e.g. Macher *et al.*, 2016). Also, extinction events and loss of biodiversity can go unnoticed. In this regard, using molecular methods to study freshwater species in mountain ranges is especially promising since many mountain ranges have been shown to harbour a large number of cryptic freshwater species (e.g. Pauls *et al.*, 2009; Katouzian *et al.*, 2016; Mamos *et al.*, 2016). Further, mountain ranges comprise many different habitats due to their topographic and climatic complexity, often leading to different species communities occurring within a small geographic area (Finn and Leroy, 2005; Múrria *et al.*, 2014; Cauvy-Fraunié *et al.*, 2015). Topography and climate form natural barriers to dispersal and many taxa occurring in mountain ranges show phenotypic and genetic adaptation to the highly differing conditions along the altitudinal gradient (Liebherr, 1986; Bonin *et al.*, 2006; Keller *et al.*, 2013; Watanabe *et al.*, 2014), ultimately leading to mountain ranges being centres of high

biodiversity. Using molecular methods to study species in such environments can help understand species diversity, ecology and their possible genetic adaptation to different habitats. Further, it can allow inferring the potential loss of species and genetic diversity when the environment changes.

Here, we analysed diversity and spatial distribution patterns in a common European stream invertebrate taxon, the freshwater limpet *Ancylus fluviatilis* (MÜLLER, 1774) sensu lato, in the Montseny mountain range on the Iberian Peninsula. The Montseny is part of the Catalan pre-coastal range (North East Iberian Peninsula). It is located at the intersection of the warm and arid climate of the mediterranean lowlands and the cooler and more precipitation rich climate of the mountainous region reaching to the Pyrenees (Thuiller *et al.*, 2003). There are three main catchments within this area, all of which are characterised by steep altitudinal gradients ranging from less than 500 to 1706 metres above sea level (masl) within approximately 10 kilometres and thus, comprising highly variable climatic conditions (Peñuelas and Boada, 2003; Jump *et al.*, 2007). We chose to study the widespread hololimnic freshwater limpet *Ancylus fluviatilis* sensu lato, because it is known to comprise several cryptic species (Hubendick, 1970; Pfenninger *et al.*, 2003; Albrecht *et al.*, 2006). Of those species, Clade 1 and Clade 4 (Pfenninger *et al.*, 2003) potentially co-occur in the North East Iberian Peninsula but have never been found in the here studied area. On a European scale, Pfenninger et al. (2003) found that the different *A. fluviatilis* species differ significantly in their ecological demands: while Clade 1 prefers cooler areas with precipitation-rich summers, Clade 4 occurs mainly in arid, generally hotter areas. Both climatic conditions can be found in the Montseny, making it an ideal area for studying the number and distribution of species within the *A. fluviatilis* species complex. *A. fluviatilis* sensu stricto is able to disperse over longer distances (Cordellier and Pfenninger, 2008), e.g. by passive transport via waterbirds and other organisms (Rees, 1965), a phenomenon commonly found in other snails and freshwater molluscs (Rees, 1965; Boag, 1986; Van Leeuwen *et al.*, 2012). The current distribution of species in the *A. fluviatilis* species complex is thus expected to be limited by ecological demands (Cordellier and Pfenninger, 2008).

In this study, we expected 1) to find morphologically cryptic species of the *A. fluviatilis* species complex in the Montseny mountain range, 2) that different *A. fluviatilis* species seldom co-occur in syntopy due to different ecological demands or interspecific competition, and 3) that species show a pattern of strong population structure within streams or catchments of the Montseny due to ecological preferences or local adaptation to bioclimatic conditions in different altitude zones. In contrast, a weaker population structure is expected between streams and catchments located in the same altitude zone, because bioclimatic conditions are more similar here and *A. fluviatilis* is known to be able to disperse over catchment boundaries, e.g. via vectors such as waterbirds.

To test these hypotheses, we first analysed the mitochondrial cytochrome c oxidase subunit 1 gene (COI) to determine the number and distribution of *Ancylus* species found in the Montseny. Second, we used a modelling approach based on bioclimatic variables to identify variables that might help to explain the occurrence of species and third, we performed population genetic analyses to investigate intraspecific partitioning of variation.

MATERIALS AND METHODS

Sampling

Sampling was performed in the Montseny mountain range (located on the North East Iberian Peninsula, Fig. 1a) and the direct surrounding area in September 2013. 44 sites were checked for the presence of *A. fluviatilis*, which was found in 36 of these sites (see Table A1 for coordinates). The three main catchments (Tordera, Besòs, Ter) and an altitudinal gradient from 120 masl to 1295 masl (Fig. 1b) were covered by the sampling. *Ancylus* specimens were collected by hand picking specimens from stones in the streams. All specimens were immediately stored in 70% ethanol, later transferred to 96% ethanol and stored at 4°C until further analysis.

A sampling permit for protected areas (Parc Natural del Montseny) was obtained from the park management prior to sampling.

135 **DNA extraction, amplification, and sequencing**

136 DNA was extracted from muscle tissue of 180 specimens (5 per site, 36 sampling sites)
137 using a salt extraction protocol (Sunnucks and Hales, 1996) (Overview of the samples: Table
138 A1). A 658 bp-fragment of the barcoding gene COI was amplified using the primers
139 LCO1490 and HCO2198 (Folmer *et al.*, 1994). The PCR mix was prepared using the
140 following protocol: 1 x PCR buffer, 0.2 mM dNTPs, 1 µl of DNA template, 0.025 U/µl
141 Hotmaster Taq (5 PRIME GmbH, Hilden, Germany) and 0.5 µM of each primer. The mix was
142 filled up to 25 µl with sterile H_2O and placed in a thermocycler for amplification. PCR settings
143 for the COI amplification were: initial denaturation at 94°C for 2 min; 36 cycles of
144 denaturation at 94°C for 20 s, annealing at 46°C for 30 s, extension at 65°C for 60 s; final
145 extension at 65°C for 5 min. 9 µl of the PCR product were purified enzymatically with 10 U of
146 Exonuclease I and 1 U Shrimp Alkaline Phosphatase (Thermo Fisher Scientific, Waltham) by
147 incubating at 37°C for 25 min and a denaturation step at 80°C for 15 min. Bidirectional
148 sequencing was performed on an ABI 3730 sequencer by GATC Biotech (Constance,
149 Germany).
150
151 **Species delimitation**
152 Raw reads were assembled and edited using Geneious 6.0.5 (Biomatters). The MAFFT
153 plugin (v. 7.017,Katoh and Standley, 2013) in Geneious was used to compute a multiple
154 sequence alignment (automatic algorithm selection, default settings). The final length of the
155 cropped alignments was 655 bp. The alignment was translated into amino acids using
156 translation table 5 (invertebrate mitochondrial codon usage table) to make sure that no stop
157 codons were present. The best model of evolution for further analyses of the data was
158 selected with jModeltest 2.1.2 (Darriba *et al.*, 2012)(default settings). Fabox (Villesen, 2007)
159 was used to collapse sequences into haplotypes. PopART (v.1, Leigh and Bryant, 2015) was
160 used to create statistical parsimony haplotype networks (Clement *et al.*, 2000) with a 95%
161 connection limit.

Two approaches were used to test for the presence of cryptic species in *Ancylus fluviatilis* sensu lato: first, the tree-based Generalized Mixed Yule Coalescent (GMYC) approach (Pons *et al.*, 2006) and second, the automated distance-based barcode gap determination approach (ABGD, Puillandre *et al.*, 2012). An ultrametric tree for all unique COI haplotypes was calculated for the GMYC analyses using BEAST v.1.8.0 (Drummond *et al.*, 2012). BEAST was run for 10 million MCMC generations, sampling every 100th tree and using both standard coalescent and the GTR + G sequence evolution model. Tracer v.1.6 (Rambaut *et al.*, 2013) was used to test for effective sampling size (ESS) and convergence of parameters. TreeAnnotator v.1.8 (Rambaut and Drummond, 2013) was used to generate a linearized consensus tree, discarding the first 3000 trees as burn-in. R v. 3.1.1 (R Core Team 2014) was used for analysis of the resulting tree with 'SPLITS' (Species Limit by Threshold Statistics) (Ezard *et al.*, 2009) with the single threshold model to test for the presence of multiple species within the dataset. The second approach used for species delimitation was ABGD. Default settings were used, with Pmax=0.1 and the K2P-model of distance correction (Kimura, 1980), as this is the common approach in DNA barcoding studies. Once the number of genetic clades within the dataset was determined, specimens from each group were blasted against the Barcode of Life database (Ratnasingham and Hebert, 2007) to verify species assignment. The *Ancylus* sequences from Pfenninger *et al.* (2003, accession numbers AY350509 - AY350525) were downloaded and aligned with the sequences generated in this study to verify assignment of sequences to one of the known cryptic species. Alignments for each species were created with Geneious and networks were computed with popArt as described above. QGIS (v 2.8, available from www.qgis.org) was used to create distribution maps.

Bioclimatic variables analyses

The bioclimatic preferences of species were modelled using a maximum entropy method in MaxEnt 3.3.3e (Phillips and Dudík, 2008), which has been shown to work well with small sample sizes (Pearson *et al.*, 2007). The region modelled was part of the North East Iberian

Peninsula (area between coordinates 42°18'N, 1°48'E , 41°06'N, 3°00'E; WGS84). A total of 19 climate layers in the 30 arc-seconds grid were obtained from WorldClim (Hijmans *et al.*, 2005) and resampled to a cell size of 800 x 800 m. WorldClim datasets are based upon standard meteorological precipitation and temperature measurements, which are transformed into bioclimatic variables (Hijmans *et al.*, 2005). These datasets are commonly used as predictor variables in species distribution modelling. To avoid using highly non-independent variables in the analyses and thus omit overfitting of models, a Spearman's rank correlation tests was performed across all pairs of variables using ENMtools (Warren *et al.*, 2010) and R (R Core Team 2015). The correlation coefficient values used as thresholds beyond which values were treated as independent were 0.7, 0.8 and 0.9. All species presence points were used to build the model; 25% of the presence points were retained for training the model. All models were run 10 times with random partitioning of training and validation points. The accuracy of all computed models was evaluated with the area under receiver operation characteristic curve which was also used to choose models for use in further analyses (Boubli and De, 2009). Range overlap and bioclimatic niche overlap were computed by using Schoener's D statistics as implemented in ENMtools. The values range from 0 (meaning no bioclimatic niche overlap) to 1 (identical range and bioclimatic niche, respectively).

Geographic partitioning of genetic variation

Φ_{ST} values as an indicator of population subdivision were calculated separately for all species found in the *A. fluviatilis* species complex using the software Arlequin (v. 3.11, Excoffier *et al.*, 2005). Φ_{ST} was chosen since it takes population history (number of mutations between haplotypes) into account. For analyses of population differentiation between altitude zones, populations of all species were classed in three groups: <500 masl, 500-1000 masl and >1000 masl as in Murria *et al*. (2014). For analyses of population differentiation between catchments, populations were classed as belonging to one of the three catchments (Tordera, Besòs, Ter) and Φ_{ST} values between groups were calculated.

The Bayesian Clustering software GENELAND (v.4.0.5 as implemented in R; Guillot *et al.*, 2005), was used to further analyse population structure in the found *Ancylus* species. Fabox was used to extract variable sites from alignments of the found species and PGDSpider (Excoffier and Lischer, 2010) was used to convert these alignments files into the GENELAND format. The settings used for running GENELAND were: Five independent runs with a maximum of 10 populations, 300 nuclei, 10 million iterations, thinning interval of 10 000, resulting in 1000 retained trees. The first 200 trees were discarded as burn-in.

RESULTS

Molecular species delimitation

Ancylus was found in 36 out of 44 sampled sites (Fig. 1b, see Table A1 for coordinates). A total of 180 specimens were analysed for the COI barcoding gene. The 655 bp alignment had 54 (9.2%) variable sites and a GC content of 29.6%. The null model of a single species was rejected both with the GMYC (likelihood ratio for single threshold model: 31.68, $p < 0.001$) and the ABGD approach (Pmax 0.1%). Both ABGD and GMYC suggested the presence of two groups in *A. fluviatilis* sensu lato. Blast searches against the Barcode of Life database assigned all sequences of both molecularly identified clades to either *A. fluviatilis* Clade 1 or *A. fluviatilis* Clade 4, both submitted by Pfenninger et al. (2003). Alignment of the generated sequences with those obtained from Genbank clustered 102 specimens with *A. fluviatilis* Clade 1, while 78 sequences clustered with *A. fluviatilis* Clade 4. Both clades were defined by Pfenninger et al. (2003). Clade 1 corresponds to *A. fluviatilis* sensu stricto, while Clade 4 is a yet undescribed species with circum- mediterranean distribution (Pfenninger et al. 2003). The species are referred to as *Ancylus* C1 and *Ancylus* C4.

Bioclimatic characterisation

For the modelling approach based on the 19 bioclimatic variables obtained from WorldClim, 3, 6 and 10 variables were retained after Spearman's Rank Correlation tests with thresholds of 0.7, 0.8 and 0.9, respectively. The 6 variable model with a Spearman's Rank Correlation

246 threshold of 0.8 resulted in good area under the receiver operating characteristic curve
247 (AUC) values for both species (*Ancylus* C1: 0.96, *Ancylus* C4: 0.89), thus this model was
248 chosen for all further analyses to mediate between lower variable correlation and higher
249 model fitting (see Table A2 for all AUC values and variables). The best explaining bioclimatic
250 variables for the occurrence of *Ancylus* C1 were the variables bio7 ("Temperature Annual
251 Range") and bio19 ("Precipitation of Coldest Quarter"). Occurrence of *Ancylus* C4 was best
252 predicted by the variables bio7 ("Temperature Annual Range") and bio15 ("Precipitation
253 Seasonality") (Table A2). Bioclimatic niche overlap for *Ancylus* C1 and *Ancylus* C4 was
254 0.643, the range overlap computed for occurrence likelihoods of >50% was 0.728 (Table A3).
255
256 **Geographic partitioning of genetic variation within *Ancylus* species**
257 Both *Ancylus* C1 and *Ancylus* C4 were found in all three studied catchments of the
258 Montseny. *Ancylus* C1 was found at 22 sampling sites and *Ancylus* C4 at 17 sampling sites.
259 Both species occurred in syntopy at 4 sampling sites (11.43%; <500 masl zone: 1 site, 500-
260 1000 masl zone: 2 sites, >1000 masl zone: 1 site; Fig. 2). *Ancylus* C1 was found more often
261 at higher altitude sites (332 – 1295 masl, median 665 masl) than *Ancylus* C4, which was
262 mainly found at lower altitude sites (120-1172 masl, median 440 masl)
263 *Ancylus* C1 showed significant population differentiation between the Tordera and Besòs
264 (Φ_{ST}: 0.322, p=0.00001) and the Tordera and Ter catchment (Φ_{ST}: 0.167, p=0.0001)(Table
265 A4). The most common haplotype (HC1_1) was found at 14 sites and in all three catchments
266 (Tordera: 5 sites, Besòs: 4 sites, Ter: 5 sites)(Fig. 3a). HC1_2 was found at five sites, of
267 which four are located in the Tordera catchment and one in the Ter catchment. HC1_3 was
268 found at 9 sites and all three catchments (Tordera: 6 sites, Besòs: 2 sites, Ter: 1 site).
269 Haplotypes C1_4, C1_5, C1_6, C1_7 and C1_8 were found in a maximum of two specimens
270 each and at single sampling sites only. Significant population differentiation in *Ancylus* C1
271 was also found between the altitude zones <500 masl and >1000 masl (Φ_{ST}: 0.343,
272 p=0.00001) and between 500-1000 masl and >1000 masl (Φ_{ST}: 0.274, p=0.0001)(Table A4).
273 GENELAND found three geographically defined groups in *Ancylus* C1. Group 1 contains the

sampling sites dominated by HC1_2, mainly lying above 1000masl (4 out of 5 sites). Group 2 contains sampling sites mainly dominated by HC1_1 (Populations in all altitude zones, but mainly (9 sites) in the 500- 1000 masl zone). Group 3 contains sampling sites mainly dominated by HC1_3 (500-1000 masl zone: 5 sites; <500masl zones: 3 sites) (Fig. 3a). A maximum of three substitutions were found between haplotypes of *Ancylus* C1 (Fig. 3b). *Ancylus* C4 showed significant population differentiation between the Tordera and Besòs (Φ_{ST}: 0.408, p=0.0001) and the Besòs and Ter catchments (Φ_{ST}: 0.876, p=0.00001), respectively. The most common haplotype (HC4_3) was found at 13 sampling sites (Besòs catchment: 9 sites, Tordera catchment: 4 sites). The second most common haplotype (HC4_1) was found in all three catchments (Tordera catchment: 4 sites, Ter catchment: 1 sites, Besòs catchment: 1 site). The haplotypes C4_2, C4_4 and C4_5 were found at single sites and in single specimens only (Fig. 3c). In *Ancylus* C4, significant population differentiation was found between populations in the <500 masl and the 500-1000 masl zone (Φ_{ST}: 0.335, p=0.009) as well as between the <500 masl and >1000 masl (Φ_{ST}: 0.667, p=0.00001) zone. GENELAND found two geographically defined groups in *Ancylus* C4. Group 1 contains four sampling sites located in the northern and eastern parts of the study area, mainly dominated by haplotype C4_1 (>1000 masl zone: 2 sites, <500 masl zone: 2 sites). Group 2 contains 13 sampling sites in the mid and western part of the Montseny, mainly dominated by haplotype C4_3 (<500 masl zone: 10 sites, 500-1000 masl zone: 3 sites)(Fig. 3c). A maximum of two substitutions were found between haplotypes of *Ancylus* C4 (Fig. 3d).

DISCUSSION

In this study, we investigated the number, distribution and genetic variation of *Ancylus fluviatilis* sensu lato species in the Montseny mountain range on the Iberian Peninsula. Our first expectation was that cryptic species of *Ancylus fluviatilis* sensu lato are present in the study area. This expectation was met by the discovery of two species occurring in the Montseny. Both *Ancylus* species were initially delimited by Pfenninger *et al*. (2003). While

302 *Ancylus* C1 corresponds to *A. fluviatilis* sensu stricto, *Ancylus* C4 is a yet undescribed
303 species with a Mediterranean distribution, ranging from Portugal through the Southern
304 Iberian Peninsula to Italy (Pfenninger *et al.*, 2003; Albrecht *et al.*, 2006). Although the
305 distribution and ecology of cryptic *A. fluviatilis* sensu lato species are roughly known on a
306 European scale, our study is the first that allows assessing the differences regarding small
307 scale distribution, population structure and bioclimatic preferences of two *A. fluviatilis* sensu
308 lato species in the same region, allowing for a better understanding of species' ecologies. In
309 the future, this might help with identification of species and improving stream quality
310 assessments by making it possible to assign correct ecological traits to species.
311 Our second expectation was that different *A. fluviatilis* sensu lato species rarely occur in
312 syntopy, possibly due to different ecological demands or competition. This expectation was
313 clearly met, as *Ancylus* C1 and *Ancylus* C4 occurred in syntopy in only 11.43% of the sites
314 and showed strong altitudinal partitioning: *Ancylus* C1 was found mainly at higher altitudes,
315 while *Ancylus* C4 was mainly found at lower elevations. As elevation is a good indicator for
316 bioclimatic and environmental conditions such as temperature (-0.65°C per 100m increase
317 in altitude on average; Dodson and Marks, 1997) and flow velocity (due to steeper mountain
318 slopes at higher altitudes), the observed pattern hints at different bioclimatic preferences of
319 the two species. This is supported by the MaxEnt modelling approach, which suggests that
320 *Ancylus* C1 occurs in areas with lower mean annual temperature and strong precipitation
321 during the cooler season of the year, while *Ancylus* C4 occurs in areas with strong
322 precipitation seasonality and higher mean annual temperature. These findings correspond to
323 those of Pfenninger et al. (2003) who, on a European scale, found *Ancylus* C1 to mainly
324 inhabit cooler, precipitation rich areas and *Ancylus* C4 to mainly inhabit hotter, seasonally dry
325 areas. Similar patterns have been observed in other aquatic invertebrate species (e.g.
326 Monaghan *et al.*, 2005; Múrria *et al.*, 2014). Closely related species co-occurring in the same
327 area often inhabit different habitats due to different bioclimatic preferences or competition
328 exclusion principle (e.g. Fišer *et al.*, 2015). In southern Europe, cooler and precipitation rich
329 conditions are mainly found at higher elevations, suggesting that *Ancylus* C1 might be close

330 to the southern border of its distribution range in the Montseny. For aquatic species, higher
331 precipitation means that more water is available during at least parts of the year, which is
332 especially important in generally dry areas or areas with high precipitation seasonality. It
333 appears thus possible that *Ancylus* C1 is more relying on constant flow of streams it inhabits,
334 while *Ancylus* C4 might be able to cope with intermittent conditions in streams and generally
335 higher water temperatures in lowland streams. This is especially important in the light of
336 future climate change and ongoing human activities such as water abstraction, which might
337 greatly alter temperature and precipitation patterns, ultimately changing flow regimes and
338 thus possibly driving some species into local extinction while giving other species the
339 possibility to colonise new habitats. Knowing the number and ecology of species in an area
340 allows tracking the impact of such changes and preventing the loss of species.
341 Third, we expected the studied species to show a pattern of intraspecific altitudinal
342 population structure due to local adaptation or ecological preferences, but not between
343 catchments due to the reported dispersal abilities of *A. fluviatilis* and the fact that streams in
344 the same altitude zone have equal bioclimatic conditions irrespective of the catchment they
345 are located in. Tests of genetic differentiation, albeit based on a limited number of specimens
346 and on the mitochondrial COI gene only, revealed that both *Ancylus* species found in the
347 Montseny show a division into lower altitude and higher altitude populations that differ
348 genetically. *Ancylus* C1 populations from above 1000 masl significantly differed from
349 populations below 1000 masl, which was also affirmed by the GENELAND results. *Ancylus*
350 C4 populations from below 500 masl differed significantly from populations located above
351 that altitude, likely due to the fact that the common haplotype C4_3 was found mainly below
352 500 masl. GENELAND, however, did not confirm the existence of a haplotype group
353 corresponding with altitude in *Ancylus* C4, probably due to the low number of specimens
354 from above 500 masl sampling sites. A pattern of genetic differentiation between higher-
355 altitude and lower-altitude populations could hint at two phenomena, possibly in combination:
356 One possibility is that *A. fluviatilis* sensu lato species are weak dispersers that rarely migrate
357 over longer distances within and between streams, thus over time populations diverge

358 genetically due to limited gene flow. Weak dispersal capabilities have been found in other
359 freshwater gastropods (Kappes and Haase, 2012). However, at least *A. fluviatilis* sensu
360 stricto has been shown to disperse over longer distances (Cordellier and Pfenninger, 2008),
361 e.g. via passive transport (Rees, 1965). The second explanation could be that populations
362 from higher and lower altitudes differ genetically due to selective processes, having adapted
363 to the different bioclimatic conditions. This corresponds to findings in other studies that found
364 high-altitude populations of species to be genetically different and potentially adapted to
365 harsher bioclimatic conditions (e.g. McCulloch *et al.*, 2009; Dussex *et al.*, 2016). Fast
366 adaptation to bioclimatic conditions has been found in *Ancylus* C1 (*A. fluviatilis* sensu
367 stricto)(Cordellier and Pfenninger, 2008), possibly making this explanation for the pattern
368 observed in the Montseny more likely. GENELAND analyses hint at *Ancylus* C1 being a
369 stronger disperser than *Ancylus* C4, as the common haplotype C1_1 was evenly found in all
370 three catchments, while a clear East West division of haplotypes was found in *Ancylus* C4.
371 Dispersal through waterbirds or mammals, as has been found in snails and other freshwater
372 taxa (Segerstråle, 1954; Figuerola and Green, 2002; Haun *et al.*, 2012; Van Leeuwen *et al.*,
373 2012), might be more common in *Ancylus* C1 than in *Ancylus* C4, possibly indicating that the
374 latter species is a weaker disperser. It can only be speculated that this might be due to
375 different dispersal strategies of the species, different preferred microhabitats that make it
376 unlikely for *Ancylus* C4 to attach to waterbirds or due to lower survival rates when being out
377 of the water. Also, it is possible that *Ancylus* C4 is not good at establishing new populations
378 when arriving in habitats already occupied by *Ancylus* C1, potentially due to lower
379 competitiveness. Further studies addressing the ecologies of both species need to be
380 conducted, ideally using a combination of laboratory and field experiments. Significant
381 population differentiation in *Ancylus* C1 between catchments in the Montseny can be
382 explained by the fact that no population containing the haplotype dominating above 1000m
383 (HC1_1) is located in the Besòs catchment and only one is found in the Ter catchment.
384 Population differentiation in *Ancylus* C4 is high between the Tordera and Besòs catchment,
385 which was also confirmed by the GENELAND results. An East West segregation of

haplotypes exists in C4, largely corresponding to the catchments Tordera and Besòs and possibly hinting at different colonisation events in the past or adaptations to bioclimatic conditions, which could not be identified in this study. The high population differentiation between Ter and Besòs catchments should be neglected due to only one sampling site from the Ter catchment being included in the analyses. Overall, it remains possible that the pattern observed is mainly due to generally low genetic variation within the species and the relatively low number of sequenced specimens. The data is based on a limited number of specimens only and needs to be interpreted with care. Studies involving nuclear markers and possibly a greater number of specimens per site are needed to verify the observed patterns of genetic variation on a geographically small scale. However, the patterns of population differentiation between altitudinal zones and catchments demonstrate the need to take population structure into account when planning to protect species and ecosystems. Environmental changes in the lower or higher altitude zones or in catchments might lead to the extinction of adapted genotypes. The resulting lower levels of genetic diversity possibly limit the adaptive potential of species and their ability to respond to environmental changes, ultimately leading to loss of genetic diversity in the species as a whole and increasing the risk of extinction (Bálint *et al.*, 2011).

Our study highlights the importance and potential of using molecular techniques to study species diversity and genetic diversity of species. Molecular studies can greatly help to understand the impact of climate change and other human stressors on biodiversity (Bálint *et al.*, 2011; Hampe and Jump, 2011; Pauls *et al.*, 2013; Macher *et al.*, 2016). Here, we found cryptic species within a common freshwater taxon and genetic divergence within species on a small geographic scale. Our results show that patterns of genetic diversity, connectivity and bioclimatic preferences can be different even between closely related species, a fact that should be considered in biomonitoring and conservation plans. Knowledge of freshwater species' diversity and ecological preferences is also important due to the fact that many bioassessment and monitoring programs worldwide rely on species occurrence data as a metric to measure ecosystem quality (e.g. Carter and Resh, 2001; Stark *et al.*, 2001; Haase

et al., 2004) and the indication value of species is mostly derived from their ecological demands. Generally, not considering molecular data and cryptic species' ecologies in monitoring programs can lead to strongly biased assessment results and ultimately to wrong management plans. The latter is especially problematic as management programmes often need to focus on protecting the maximum amount of biodiversity with the least amount of monetary effort. Using molecular methods can help to identify and effectively protect species and intraspecific diversity.

ACKNOWLEDGMENTS

We thank Nuria Bonada for help with acquiring the sampling permission for the Parque Natural Montseny and helpful discussions. The Parque Natural Montseny is thanked for giving us permission to collect samples in the park. We thank Alexander M. Weigand for helpful discussions, Lisa Poettker for help with sampling and lab work and Raúl González Pech for translating the abstract and figure captions. We are indebted to the North-Rhine Westphalian Academy of Sciences for financial support.

DATA ACCESSABILITY

CO1 DNA sequences:

GenBank accession numbers for *Ancylus* Clade 1: XXX-XXX. Will be available upon publication and numbers will be added accordingly.

GenBank accession numbers for *Ancylus* Clade 4: XXX-XXX. Will be available upon publication and numbers will be added accordingly.

CONFLICT OF INTEREST

None declared

REFERENCES

Adams, M., Raadik, T., Burridge, C. and Georges, A., 2014. Global biodiversity assessment and hyper-cryptic species complexes: more than one species of elephant in the room? Syst. Biol. 63, 518–533.

Albrecht, C., Trajanovski, S., Kuhn, K., Streit, B. and Wilke, T., 2006. Rapid evolution of an ancient lake species flock: freshwater limpets (Gastropoda: Ancylidae) in the Balkan Lake Ohrid. Org. Divers. Evol. 6, 294–307.

Amato, A., Kooistra, W., Ghiron, J., Mann, D., Pröschold, T. and Montresor, M., 2007. Reproductive isolation among sympatric cryptic species in marine diatoms. Protist 158, 193–207.

Bálint, M., Domisch, S., Engelhardt, C., Haase, P., Lehrian, S., Sauer, J., Theissinger, K., Pauls, S. and Nowak, C., 2011. Cryptic biodiversity loss linked to global climate change. Nat. Clim Change 1, 313–318.

Boag, D., 1986. Dispersal in pond snails: potential role of waterfowl. Can. J. Zool. 64, 904–909.

Bonin, A., Taberlet, P., Miaud, C. and Pompanon, F., 2006. Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (Rana temporaria). Mol. Biol. Evol. 23, 773–783.

Boubli, J. and De, L., 2009. Modeling the geographical distribution and fundamental niches of Cacajao spp. and Chiropotes israelita in Northwestern Amazonia via a maximum entropy algorithm. Int. J. Primatol. 30, 217–228.

Carter, J. and Resh, V., 2001. After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. J N Am Benthol Soc 20, 658–682.

Cauvy-Fraunié, S., Espinosa, R., Andino, P., Jacobsen, D. and Dangles, O., 2015. Invertebrate metacommunity structure and dynamics in an andean glacial stream network facing climate change. PLoS ONE 10, e0136793.

Clement, M., Posada, D. and Crandall, K., 2000. TCS: a computer program to estimate gene genealogies. Mol Ecol 9, 1657–1659.

Cordellier, M. and Pfenninger, M., 2008. Climate□driven range dynamics of the freshwater limpet, Ancylus fluviatilis (Pulmonata, Basommatophora). J Biogeogr 35, 1580–1592.

Darriba, D., Taboada, G., Doallo, R. and Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9, 772.

Dodson, R. and Marks, D., 1997. Daily air temperature interpolated at high spatial resolution over a large mountainous region. Clim. Res. 8, 1-20.

Drummond, A., Suchard, M., Xie, D. and Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29, 1969–1973.

Dussex, N., Chuah, A. and Waters, J., 2016. Genome□wide SNPs reveal fine□scale differentiation among wingless alpine stonefly populations and introgression between winged and wingless forms. Evolution 70, 38–47.

Excoffier, L., Laval, G. and Schneider, S., 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1, 47–50.

Excoffier, L. and Lischer, H., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10, 564–567.

Ezard, T., Fujisawa, T. and Barraclough, T., 2009. SPLITS: species' limits by threshold statistics. R Package Version 1.

Figuerola, J. and Green, A., 2002. Dispersal of aquatic organisms by waterbirds a review of past research and priorities for future studies. Freshw Biol 47, 483–494.

Finn, D. and Poff, L., 2005. Variability and convergence in benthic communities along the longitudinal gradients of four physically similar Rocky Mountain streams. Freshw Biol 50, 243–261.

Fišer, Z., Altermatt, F., Zakšek, V., Knapič, T. and Fišer, C., 2015. Morphologically Cryptic Amphipod Species Are "Ecological Clones" at Regional but Not at Local Scale: A Case Study of Four Niphargus Species. PLoS ONE 10, e0134384.

Folmer, O., Black, M., Hoeh, W., Lutz, R. and Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3, 294–299.

Guillot, G., Mortier, F. and Estoup, A., 2005. Geneland: a computer package for landscape genetics. Mol Ecol Notes 5, 712–715.

Haase, P., Lohse, S., Pauls, S., Schindehütte, K., Sundermann, A., Rolauffs, P. and Hering, D., 2004. Assessing streams in Germany with benthic invertebrates: development of a practical standardised protocol for macroinvertebrate sampling and sorting. Limnol.-Ecol. Manag. Inland Waters 34, 349–365.

Hampe, A. and Jump, A., 2011. Climate relicts: past, present, future. Annu. Rev. Ecol. Evol. Syst. 42, 313–333.

Haun, T., Salinger, M., Pachzelt, A. and Pfenninger, M., 2012. On the processes shaping small-scale population structure in Radix balthica (Linnaeus 1758). Malacologia 55, 219–233.

Hijmans, R., Cameron, S., Parra, J., Jones, P. and Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978.

Hubendick, B., 1970. Studies on Ancylidae: The palearctic and oriental species and formgroups.

Jump, A., Hunt, J. and Penuelas, J., 2007. Climate relationships of growth and establishment across the altitudinal range of Fagus sylvatica in the Montseny Mountains, northeast Spain. Ecoscience 14, 507–518.

Kappes, H. and Haase, P., 2012. Slow, but steady: dispersal of freshwater molluscs. Aquat. Sci. 74, 1.

Katoh, K. and Standley, D., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30, 772–780.

Katouzian, A., Sari, A., Macher, J., Weiss, M., Saboori, A., Leese, F. and Weigand, A., 2016. Drastic underestimation of amphipod biodiversity in the endangered Irano-Anatolian and Caucasus biodiversity hotspots. Sci. Rep. 6, 22507.

Keller, I., Alexander, J., Holderegger, R. and Edwards, P., 2013. Widespread phenotypic and genetic divergence along altitudinal gradients in animals. J Evol Biol 26, 2527–2543.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16, 111–120.

Lagrue, C., Wattier, R., Galipaud, M., Gauthey, Z., Rullmann, J., Dubreuil, C., Rigaud, T. and Bollache, L., 2014. Confrontation of cryptic diversity and mate discrimination within Gammarus pulex and Gammarus fossarum species complexes. Freshw Biol 59, 2555–2570.

Van Leeuwen, C., Velde, G., Groenendael, J. and Klaassen, M., 2012. Gut travellers: internal dispersal of aquatic organisms by waterfowl. J. Biogeogr. 39, 2031–2040.

Leigh, J. and Bryant, D., 2015. popart: full□feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110–1116.

Liebherr, J., 1986. Comparison of genetic variation in two carabid beetles (Coleoptera) of differing vagility. Ann. Entomol. Soc. Am. 79, 424–433.

Macher, J., Salis, R., Blakemore, K., Tollrian, R., Matthaei, C. and Leese, F., 2016. Multiple-stressor effects on stream invertebrates: DNA barcoding reveals contrasting responses of cryptic mayfly species. Ecol. Indic. 61, 159–169.

Mamos, T., Wattier, R., Burzyński, A. and Grabowski, M., 2016. The legacy of a vanished sea: a high level of diversification within a European freshwater amphipod species complex driven by 15 My of Paratethys regression. Mol Ecol.

McCulloch, G., Wallis, G. and Waters, J., 2009. Do insects lose flight before they lose their wings? Population genetic structure in subalpine stoneflies. Mol Ecol 18, 4073–4087.

Monaghan, M., Robinson, C., Spaak, P. and Ward, J., 2005. Macroinvertebrate diversity in fragmented Alpine streams: implications for freshwater conservation. Aquat. Sci. 67, 454–464.

Múrria, C., Morante, M., Rieradevall, A., Ribera, A. and Prat, N, 2014. Genetic diversity and species richness patterns in Baetidae (Ephemeroptera) in the Montseny Mountain range (North-East Iberian Peninsula). Limnetica 33, 313–326.

Ortells, R., Gómez, A. and Serra, M., 2003. Coexistence of cryptic rotifer species: ecological and genetic characterisation of Brachionus plicatilis. Freshw Biol 48, 2194–2202.

Pauls, S., Blahnik, R., Zhou, X., Wardwell, C. and Holzenthal, R., 2010. DNA barcode data confirm new species and reveal cryptic diversity in Chilean Smicridea (Smicridea) (Trichoptera:Hydropsychidae). J N Am Benthol Soc 29, 1058–1074.

Pauls, S., Nowak, C., Bálint, M. and Pfenninger, M., 2013. The impact of global climate change on genetic diversity within populations and species. Mol Ecol 22, 925–946.

Pauls, S., Theissinger, K., Ujvarosi, L., Balint, M. and Haase, P., 2009. Patterns of population structure in two closely related, partially sympatric caddisflies in Eastern Europe:

606 historic introgression, limited dispersal, and cryptic diversity 1. J N Am Benthol Soc
607 28, 517–536.
608
609 Pearson, R., Raxworthy, C., Nakamura, M. and Townsend, P., 2007. Predicting species
610 distributions from small numbers of occurrence records: a test case using cryptic
611 geckos in Madagascar. J. Biogeogr. 34, 102–117.
612
613 Peñuelas, J. and Boada, M., 2003. A global change□induced biome shift in the Montseny
614 mountains (NE Spain). Glob. Change Biol. 9, 131–140.
615
616 Pfenninger, M. and Schwenk, K., 2007. Cryptic animal species are homogeneously
617 distributed among taxa and biogeographical regions. BMC Evol Biol 7, 121.
618
619 Pfenninger, M., Staubach, S., Albrecht, C., Streit, B. and Schwenk, K., 2003. Ecological and
620 morphological differentiation among cryptic evolutionary lineages in freshwater
621 limpets of the nominal form□group Ancylus fluviatilis (O.F. Müller, 1774). Mol Ecol 12,
622 2731–2745.
623
624 Phillips, S. and Dudík, M., 2008. Modeling of species distributions with Maxent: new
625 extensions and a comprehensive evaluation. Ecography 31, 161–175.
626
627 Pons, J., Barraclough, T., Gomez-Zurita, J., Cardoso, A., Duran, D., Hazell, S., Kamoun, S.,
628 Sumlin, W. and Vogler, A., 2006. Sequence-based species delimitation for the DNA
629 taxonomy of undescribed insects. Syst Biol 55, 595–609.
630
631 Puillandre, N., Lambert, A., Brouillet, S. and Achaz, G., 2012. ABGD, Automatic Barcode
632 Gap Discovery for primary species delimitation. Mol Ecol 21, 1864–1877.
633
634 Rambaut, A. and Drummond, A., 2013. TreeAnnotator v1. 7.0.
635
636 Rambaut, A., Suchard, M., Xie, D. and Drummond, A., 2013. Tracer v1.5, Available from
637 http://beast.bio.ed.ac.uk/Tracer.
638
639 Ratnasingham, S. and Hebert, P., 2007. BOLD: the Barcode of Life Data System
640 (http://www.barcodinglife.org). Mol Ecol Notes 7, 355–364.
641
642 Rees, W., 1965. The aerial dispersal of Mollusca. J. Molluscan Stud. 36, 269–282.
643
644 Rissler, L. and Apodaca, J., 2007. Adding more ecology into species delimitation: ecological
645 niche models and phylogeography help define cryptic species in the black
646 salamander (Aneides flavipunctatus). Syst. Biol. 56, 924–942.
647
648 Segerstråle, S., 1954. The Freshwater Amphipods, Gammarus Pulex (L.) and Gammarus
649 Lacustris GO Sars, in Denmark and Fennoscandia-a Contribution to the Late-and
650 Postglacial Immigration History of the Aquatic Fauna of Northern Europe.
651
652 Stark, J. D., 2001. Protocols for sampling macroinvertebrates in wadeable streams.
653 Cawthron Institute, New Zealand
654
655 Steffen, W., Richardson, K., Rockström, J., Cornell, S., Fetzer, I., Bennett, E., Biggs, R.,
656 Carpenter, S., de Vries, W. and de Wit, C., 2015. Planetary boundaries: Guiding
657 human development on a changing planet. Science 347, 1259855.
658
659 Sunnucks, P. and Hales, D., 1996. Numerous transposed sequences of mitochondrial
660 cytochrome oxidase I-II in aphids of the genus Sitobion (Hemiptera: Aphididae). Mol
661 Biol Evol 13, 510–524.

Thuiller, W., Vayreda, J., Pino, J., Sabate, S., Lavorel, S. and Gracia, C., 2003. Large□scale environmental correlates of forest tree distributions in Catalonia (NE Spain). Glob. Ecol. Biogeogr. 12, 313–325.

Villesen, P., 2007. FaBox: an online toolbox for fasta sequences. Mol Ecol Notes 7, 965–968.

Vörösmarty, C., McIntyre, P., Gessner, M., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S., Sullivan, C., Reidy Liermann, C. and Davies, P., 2010. Global threats to human water security and river biodiversity. Nature 467, 555–561.

Warren, D., Glor, R. and Turelli, M., 2010. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33, 607–611.

Watanabe, K., Kazama, S., Omura, T. and Monaghan, M., 2014. Adaptive Genetic Divergence along Narrow Environmental Gradients in Four Stream Insects. PLoS ONE 9, e93055.

Weigand, A., Jochum, A., Pfenninger, M.m Steinke, D. and Klussmann□Kolb, A., 2011. A new approach to an old conundrum—DNA barcoding sheds new light on phenotypic plasticity and morphological stasis in microsnails (Gastropoda, Pulmonata, Carychiidae). Mol Ecol Resour 11, 255–265.

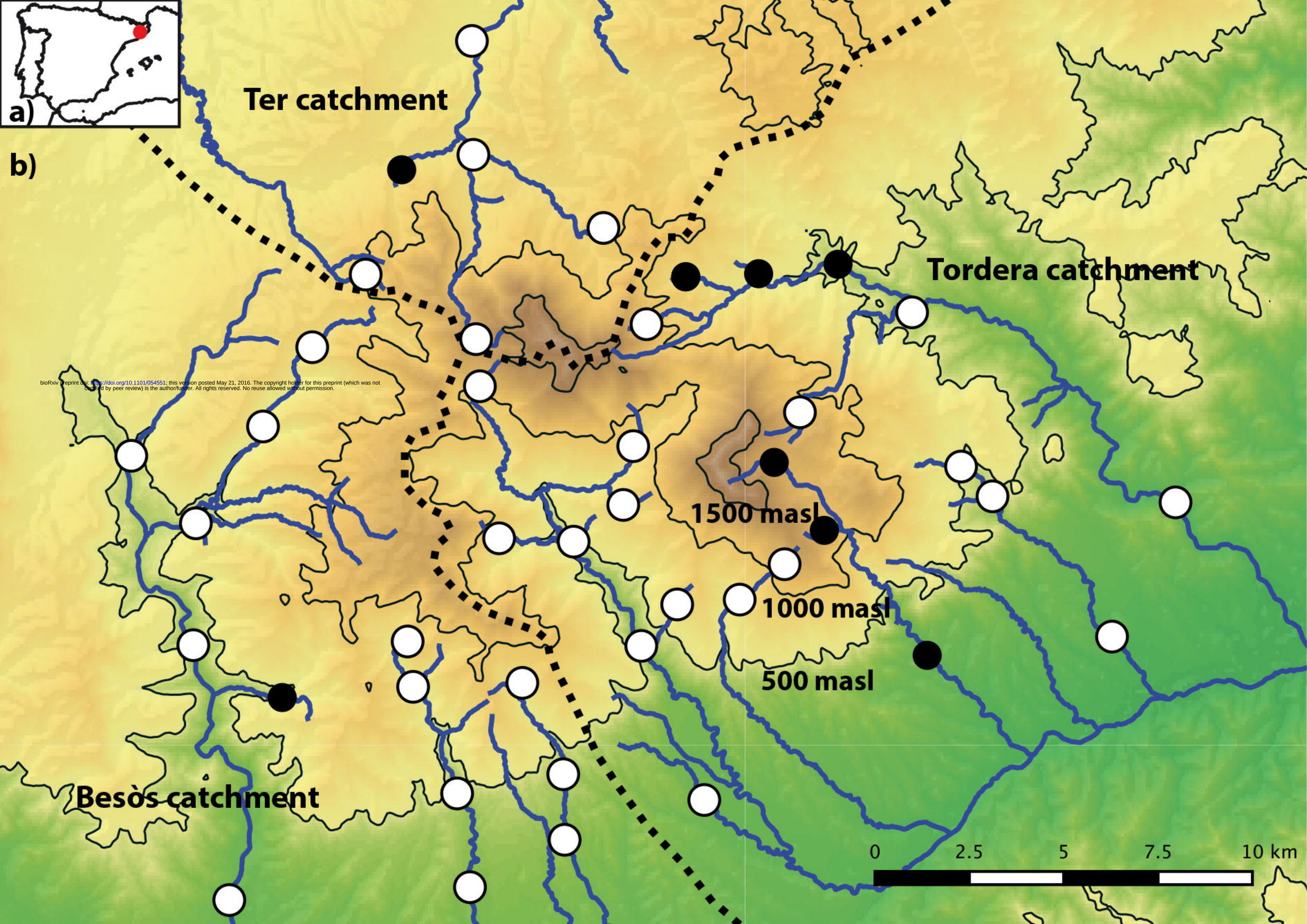
Weiss, M., Macher, J., Seefeldt, M. and Leese, F., 2014. Molecular evidence for further overlooked species within the Gammarus fossarum complex (Crustacea: Amphipoda). Hydrobiologia 721, 165–184.

FIGURE CAPTIONS

Figure 1: a) Location of the Montseny mountain range on the Iberian Peninsula (red circle); b) Location of the sampling sites in the Montseny. White dots: Sites where *Ancylus fluviatilis* sensu lato was found. Black dots: Sites where *A. fluviatilis* sensu lato was not found. Catchment boundaries are shown as dashed black lines. Rivers are shown as blue lines.

Figure 2: Map showing the occurrence of *Ancylus* Clade 1 and Clade 4 in the Montseny. Red dots indicate presence of *Ancylus* C1, blue dots presence of *Ancylus* C4. Mixed blue and red dots indicate the presence of both species at one sampling site. Black dots indicate absence of *Ancylus*. Catchment boundaries are shown as dashed black lines. Rivers are shown as blue lines.

703 Figure 3: a)+ c) Map of the study area indicating the presence of *Ancylus* Clade 1 and
704 *Ancylus* Clade 4, respectively, the number of studied specimens per sampling site and found
705 COI haplotypes. Number of specimens per sampling sites is 5 unless otherwise stated.
706 GENELAND groups are shown as coloured dashed lines and catchment boundaries as black
707 dashed lines. b)+ d) Statistical Parsimony Network of *Ancylus* Clade 1 and *Ancylus* Clade 4
708 haplotypes, respectively. Dots represent sampled haplotypes, bars represent number of
709 substitutions between haplotypes.



Ancylus C1

Ancylus C4

Ancylus C1/
Ancylus C4

Ter catchment

Tordera catchment

1500 masl

1000 masl

500 masl

Besòs catchment

0 2.5 5 7.5 10 km

a)

Ter catchment

n=2

Besòs catchment

n=1

HC1_1
HC1_2
HC1_3
HC1_4
HC1_5
HC1_6
HC1_7
HC1_8
Group C1_1
Group C1_2
Group C1_3

1500 masl

1000 masl

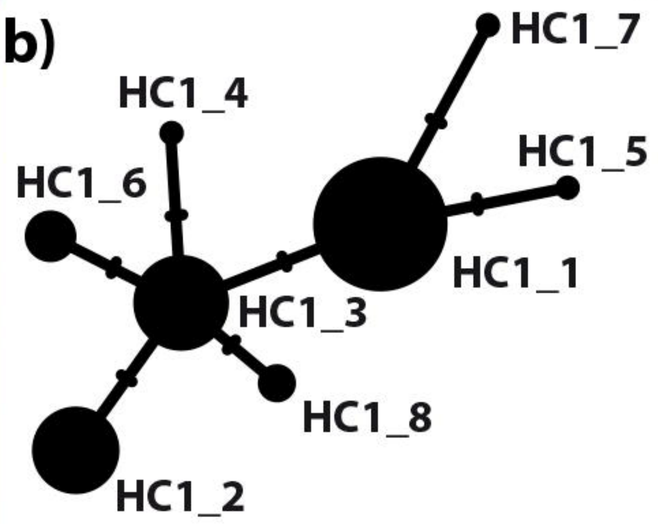
500 masl

n=1

Tordera catchment

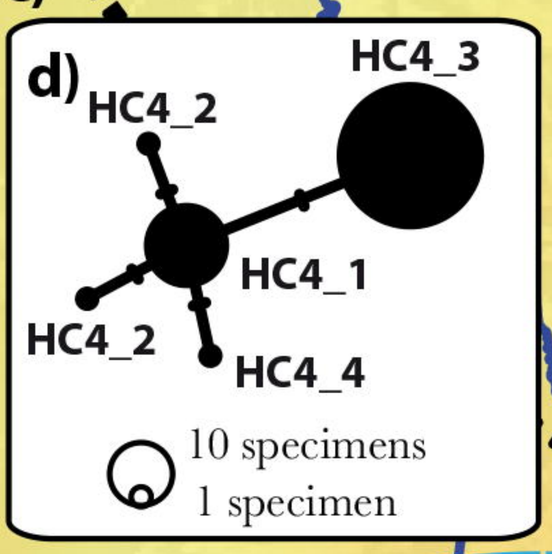
0 2.5 5 7.5 10 km

b)



c)

d)



Ter catchment

n=3

n=4

n=1

HC4_1
HC4_2
HC4_3
HC4_4
HC4_5
Group C4_1
Group C4_2

1500 masl

1000 masl

500 masl

Besòs catchment

Tordera catchment

0 2.5 5 7.5 10 km