



# **We are ready for faunistic surveys of bdelloid rotifers through DNA barcoding: the example of *Sphagnum* bogs of the Swiss Jura Mountains**

Diego Fontaneto<sup>1,\*</sup>, Ester M. Eckert<sup>1</sup>, Nikoleta Anicic<sup>1</sup>, Enrique Lara<sup>2,3</sup> and Edward A. D. Mitchell<sup>2,4</sup>

<sup>1</sup> National Research Council of Italy, Water Research Institute, Microbial Ecology Group, largo Tonolli 50, I-28922 Verbania Pallanza, Italy.

<sup>2</sup> Laboratory of Soil Biodiversity, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, CH-2000 Neuchâtel, Switzerland.

<sup>3</sup> Real Jardín Botánico, CSIC, Plaza Murillo 2, ES-28014 Madrid, Spain.

<sup>4</sup> Jardin Botanique de Neuchâtel, Chemin du Perthuis-du-Sault 58, CH-2000 Neuchâtel, Switzerland.

\* Corresponding author: diego.fontaneto@cnr.it

Received: 12/03/18

Accepted: 14/05/18

## **ABSTRACT**

**We are ready for faunistic surveys of bdelloid rotifers through DNA barcoding: the example of *Sphagnum* bogs of the Swiss Jura Mountains**

The identification of biological diversity through DNA barcoding and metabarcoding of the organisms living in the field has the potential to revolutionise the way biological surveys and monitoring are performed. Yet, we still do not know if the current representativeness of the reference database of DNA sequence data is sufficient to allow such approaches. Here, we show that, at least for bdelloid rotifers (Metazoa; Rotifera; Bdelloidea) in Europe, current knowledge is ripe to perform such surveys. We show the results of an exercise performed on bdelloid rotifers in *Sphagnum* bogs of the Swiss Jura Mountain. The results of DNA-based identifications were rather consistent with the morphology-based identifications, and the few cases of mismatch could be used as a cautionary tale to avoid potential misinterpretations of results. The mismatches were due to cases of the closest match not being genetically very close, and to the occurrence of cryptic species.

**Key words:** Bdelloidea, biodiversity, cytochrome c oxidase subunit I, Rotifera, taxonomy

## **RESUMEN**

**Listos para revisiones faunísticas de rotíferos bdelloideos por DNA barcoding: el caso de las turberas de *Sphagnum* de las montañas del Jura suizo**

*La identificación de la diversidad biológica a través de DNA barcoding y metabarcoding de los organismos en el medio ambiente tiene el potencial de revolucionar la forma en que se realizan los inventarios biológicos y el monitoreo. Sin embargo, todavía no se sabe si las bases de datos genéticos de referencia a disposición hoy en día son lo suficientemente representativas como para permitir tales enfoques. Aquí, mostramos que, al menos para los rotíferos bdelloideos (Metazoa; Rotifera; Bdelloidea) de Europa, el nivel de conocimiento es suficiente para realizar tales estudios. Mostramos los resultados de un ejercicio realizado sobre rotíferos bdelloideos en turberas de *Sphagnum* del Jura suizo. Los resultados de las identificaciones basadas en el ADN fueron bastante consistentes con las identificaciones basadas en la morfología, y los pocos casos de desajuste podrían utilizarse como una advertencia para evitar posibles interpretaciones erróneas de los resultados. Estos desajustes se debieron a que las secuencias más cercanas seguían alejadas de los organismos realmente encontrados y a la presencia de especies crípticas.*

**Palabras clave:** Bdelloidea; biodiversidad, subunidad I de la citocromo c oxidasa, Rotifera, taxonomía

## INTRODUCTION

Rotifers are one of the most common and abundant groups of animals living in continental waters (Fontaneto & De Smet, 2015). The known global richness of this phylum is not very high, with only slightly more than 2000 species described (Segers, 2007); on the other hand, local richness can be quite high, with more than 100 species occurring in a single temperate lake (Dumont & Segers, 1996; Segers & De Smet 2008). The geographic distributions of species are very wide, allowing comparisons of communities in similar ecosystems across continents (Fontaneto *et al.*, 2012). Given their ubiquity and abundance, rotifers have been suggested as useful biomonitors of environmental quality (Sládeček, 1983; Obertegger *et al.*, 2011; Kuczyńska-Kippen, 2018). Yet, their routine identification is hampered by a high degree of phenotypic plasticity in several morphological features (Gilbert, 2017) coupled with a high degree of morphological stasis for other features (Campillo *et al.*, 2005). Such taxonomic uncertainty is mirrored in the high degree of cryptic species found to date in all groups for which DNA sequences are available (García-Morales & Elías-Gutiérrez, 2013; Mills *et al.*, 2017; Kordbacheh *et al.*, 2017), with the further complication of between-species hybridisation (Suatoni *et al.*, 2006; Papakostas *et al.*, 2016; Obertegger *et al.*, 2018). Finally, as morphological identification often requires observing living specimens to see the necessary identification criteria, especially for bdelloid rotifers (Donner, 1965), samples cannot be fixed. Using DNA extracted from environmental samples (eDNA) could be a useful alternative solution for practical applications of rotifers as bioindicators, but this approach has not yet been validated.

Thanks to several studies on DNA taxonomy and population genetics in rotifers, many DNA sequences are nowadays available for several species: a GenBank search performed on February 16<sup>th</sup> 2018 gave an astounding figure of almost 10 000 hits in Rotifera for cytochrome c oxidase I (COI), the most commonly used barcoding marker to date for animals (Hebert *et al.*, 2003). Thus, such a marker could be used to obtain information

on the occurrence of species in the field, bypassing the need for a morphological approach for species identification by directly obtaining DNA sequence data to be compared with a reference database. Such an approach is what several researchers are advocating for the future of biological monitoring (Leese *et al.*, 2018), called biomonitoring 2.0 (Baird *et al.*, 2012). The identification of single organisms through DNA sequences from a reference marker is an established approach, called DNA barcoding (Hebert *et al.*, 2003). Its extension at the community level (the identification of the whole group of organisms living in a sample) is called DNA metabarcoding (Taberlet *et al.*, 2012), and is considered at the forefront of biomonitoring 2.0 (Leese *et al.*, 2018).

One of the problems of applying such an approach in the field for routine biological monitoring and faunistic studies is that we still do not know if a DNA-based identification would be applicable to rotifers, because of the presence of cryptic species and of how far the existing reference database would be exhaustive enough to provide accurate taxonomic assignments. For well-studied animals such as fish (Díaz *et al.*, 2016), Lepidoptera (Huemmer *et al.*, 2014), Ephemeroptera, Plecoptera, and Trichoptera (Morinière *et al.*, 2017), the reference databases seem reliable, but for understudied microscopic animals the reliability needs to be demonstrated. The aim of this study is to provide an example of such an application of DNA barcoding, testing its efficiency and assessing its potential biases. In order to produce a reliable test of a faunistic study performed on DNA sequence data, we focused on one group of rotifers, the bdelloids, for which taxonomic uncertainties are high (Fontaneto *et al.*, 2009) and faunistic studies are scarce. Moreover, we performed field sampling in Switzerland, where rotifer diversity is highly understudied: no records of bdelloids are available for the country in the Fauna Europaea database (de Yong *et al.*, 2014). Therein, we focused on a highly specific and understudied habitat, *Sphagnum* bogs, because the diversity of bdelloids is known to be relatively high in small water bodies with acidic waters (Donner, 1965).

The rationale of the tests included a comparison between morphological identification and

**Table 1.** List of samples analysed during the project. Acronyms for localities are: NE, Neuchâtel; JU, Jura; BE, Bern). Coordinates are in the WGS84 reference system. *Lista de muestras analizadas durante este proyecto. Los acrónimos para las localidades son: NE, Neuchâtel, JU, Jura; BE, Bern). Las coordenadas están en el sistema de referencia WGS84.*

Sample code	Locality	Description	Date	Habitat	North	East
D01	Le Cachot (NE)	Fossé Pochon (former peat extraction ditch)	07/10/14	wet <i>Sphagnum cf. fallax</i>	47.004665°	6.664461°
D02	Bois des Lattes (JU)	Pool in a secondary (cutover) part of bog	07/10/14	wet/submerged <i>Sphagnum</i> spp.	46.973649°	6.708459°
D03	Le Cachot (NE)	Large <i>Sphagnum cuspidatum</i> pool in the centre of the bog	07/10/14	wet/submerged <i>Sphagnum cuspidatum</i>	47.005440°	6.665511°
D04	Bois des Lattes (NE)	Blocked drainage ditch with <i>Carex rostrata</i> , <i>Sphagnum cf. fallax</i>	07/10/14	wet/submerged <i>Sphagnum cf. fallax</i>	46.972924°	6.707294°
D05	Le Cachot (NE)	Side of a large <i>Sphagnum cuspidatum</i> pool in the centre of the bog	07/10/14	<i>Sphagnum magellanicum</i>	47.005440°	6.665511°
D06	Le Cachot (NE)	Lawn with <i>Betula nana</i>	07/10/14	<i>Sphagnum capillifolium</i> & <i>S. magellanicum</i>	47.005738°	6.663991°
D07	Neuchâtel (NE)	Botanical Garden	11/10/14	Lower pool, with <i>Lemna</i>	46.999833°	6.934687°
D08	Neuchâtel (NE)	Botanical Garden	11/10/14	Lake, macrophytes with <i>Asellus</i>	46.999825°	6.936297°
D09	Neuchâtel (NE)	Botanical Garden	11/10/14	Higher pool, detritus	46.999959°	6.934659°
D10	Etang de la Gruère (JU)	Lake shore	13/10/14	Floating vegetation on shore: <i>Sphagnum cf. fallax</i>	47.238765°	7.052270°
D11	Etang de la Gruère (JU)	Lake shore opposite to the "peninsula"	13/10/14	Submerged vegetation	47.237978°	7.048248°
D12	Chaux des Breuleux (JU)	Floating vegetation around a large pond	13/10/14	wet <i>Sphagnum cf. fallax</i>	47.226242°	7.045749°
D13	Les Pontins (BE)	Small pool at the base of a peat slope	13/10/14	wet <i>Sphagnum cf. fallax</i>	47.127285°	6.989863°

DNA-based identification, using (1) a taxonomically informed species assignment with a reference database, and (2) an unsupervised assignment based on barcoding thresholds only. The main aim of the tests was to assess if we are now ready for such kind of inventories of biological diversity based directly on DNA and not only through morphological identification.

## MATERIAL AND METHODS

### Sampling

Authorisations were requested and obtained from local governments to sample *Sphagnum* mosses from selected peatlands in Switzerland, in the Jura Mountains (cantons of Neuchâtel, Jura and Bern) in relation to this project. Sampling took place mostly in October 2014. We collected samples in L'Etang de la Gruère (Jura/Bern), La Chaux-des-Breuleux (Bern/Jura), Les Pontins (Bern), Le Cachot (Neuchâtel) and Le Bois-des-Lattes (Neuchâtel) (Table 1). Each sample consisted in a cube of *Sphagnum* moss of 5 cm side, stored in a plastic bottle, and kept refrigerated in the field and during the transport to the laboratory.

### Species identification

Animals were sorted and isolated in the laboratory under a dissecting microscope, taking as a representative subsample a cube of 1cm side from each sample. All isolated individuals were identified to species level or to genus level. Pictures were taken at a compound microscope at 200 to 400x magnification for each isolated individual. The identification characters for bdelloids are only visible on active individuals (Donner, 1965) and it is therefore impossible to fix the animals in such a way that characters are visible on a permanent slide, whereas they can still be visible on photographs.

### DNA sequence data

DNA was extracted from single identified and photographed individuals of bdelloid rotifers using a Chelex extraction protocol (Gómez *et al.*,

2002). For each individual, partial COI mtDNA gene was sequenced adapting the protocol for monogonont rotifers (Gómez *et al.*, 2002): DNA from each single animal was extracted in 35 µL of Chelex (InstaGene Matrix; Bio-Rad, CA, USA). A 658 base pairs fragment of the COI gene was PCR amplified using optimized primers LCOI (5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3') and HCOI (5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3') (Folmer *et al.*, 1994). Cycle conditions comprised initial denaturation at 94 °C for 5 min, followed by 35 cycles of 94 °C for 1 min, 43 °C for 1 min and 72 °C for 90 s, and a final extension step of 72 °C for 7 min. Purification and sequencing were performed by an external company. Chromatograms were checked for ambiguous positions using FINCHTV 1.4.0, aligned with MAFFT (Katoh & Standley, 2013) with the default automatic settings, and visually checked by eye for correct protein coding in Mesquite (Maddison & Maddison, 2018).

### Analyses

The first test was performed on the DNA barcoding pipeline that would be used to try to identify an organism starting from its DNA sequence and comparing it to a reference database. We performed this step by checking the highest similarity in the known DNA sequences available in the GenBank database (i.e. "best match") through BLAST searches (Benson *et al.*, 2013). For each of the sequences we obtained from the animals collected in the field, we gathered information on the GenBank best match regarding species identification, percentage similarity, and the length in base pairs of the overlapping part of the sequences.

We then checked how many of the retrieved best matches correctly identified the sequence to species and genus level, in accordance with our morphological identification. Ideally, the match should be 100%; yet, biological diversity is much higher than what we can actually describe, and some level of uncertainty is always expected. For example, it could be that different morphological species match to the same species name in the GenBank reference database, or that different individuals of the same morphospecies match to

**Table 2.** List of the 17 morphologically identified species of bdelloid rotifers from *Sphagnum* peatlands in the Swiss Jura Mountains, number of sequenced individuals and from how many populations they come, number of available COI sequences in GenBank for each species, number of successfully identified individuals from BLAST searches at the species, genus, and family level. In addition, the number of ABGD units for each species is provided in the last column. The last row summarises the totals for each column. *Lista de las 17 especies de rotíferos bdelloideos identificados morfológicamente de las turberas de Sphagnum de las montañas del Jura suizo, número de individuos secuenciados y de cuántas poblaciones provienen, número de secuencias de COI disponibles en GenBank para cada especie, número de individuos identificados con éxito usando BLAST a niveles de especie, género y familia. Asimismo, se proporciona el número de unidades ABGD para cada especie en la última columna. La última fila contiene las sumas de cada columna.*

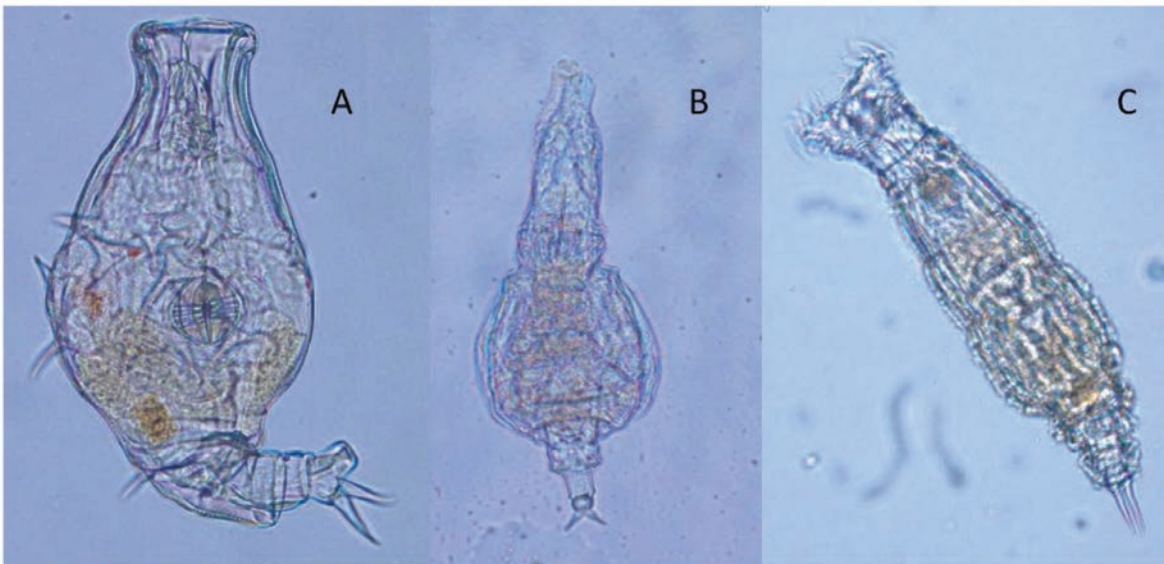
species	individuals	populations	GenBank sequences	correct species	correct genus	correct family	ABGD units
<i>Adineta gracilis</i> Janson, 1893	8	5	26	8	8	8	1
<i>Adineta steineri</i> Bartos, 1951	1	1	13	1	1	1	1
<i>Adineta vaga</i> (Davis, 1873)	1	1	230	1	1	1	1
<i>Dissotrocha aculeata</i> (Ehrenberg, 1832)	2	1	23	2	2	2	1
<i>Dissotrocha macrostyla</i> (Ehrenberg, 1838)	14	4	26	3	13	14	3
<i>Habrotrocha lata</i> (Bryce, 1892)	3	2	2	2	2	2	2
<i>Macrotrachela plicata</i> (Bryce, 1892)	2	2	3	0	1	2	1
<i>Macrotrachela quadricornifera</i> Milne, 1886	7	3	102	7	7	7	2
<i>Otostephanos donneri</i> Bartos, 1959	1	1	4	0	0	0	1
<i>Philodina citrina</i> Ehrenberg, 1832	7	3	56	7	7	7	2
<i>Philodina megalotrocha</i> Ehrenberg, 1832	1	1	26	1	1	1	1
<i>Philodina</i> sp.	1	1	NA	NA	1	1	1
<i>Rotaria macroceros</i> (Gosse, 1851)	2	1	2	0	1	1	2
<i>Rotaria magnacalcarata</i> (Parsons, 1892)	4	1	56	4	4	4	1
<i>Rotaria rotatoria</i> (Pallas, 1766)	48	6	857	48	48	48	8
<i>Rotaria</i> sp.	6	4	NA	NA	6	6	2
<i>Rotaria tardigrada</i> (Ehrenberg, 1832)	8	4	15	8	8	8	1
total	116	13	1441	92	111	113	31

different species names. Such mismatches could be due to actual taxonomical uncertainties in the identification of cryptic or pseudocryptic species based on morphology, or to potential errors in the reference database; moreover, it could also be that the closest match in the reference database is still not so genetically similar and thus provides an unreliable and false best match.

To address these questions by assessing the type of errors that produced wrong taxonomic assignments through BLAST searches and including also other potential confounding factors, we performed statistical analyses through analyses of variance (ANOVA). We tested whether the percentage of similarity, or the length of the sequence, was different between correctly and incorrectly identified sequences. Moreover, by using species-level summary data (Table 2), we tested whether the proportion of correctly identified sequences for each morphological species could be due to the number of individuals for each species, as a proxy for sampling bias in the data, or to the number of sequences available in GenBank, as a proxy for bias in the reference database. We addressed this issue by using gener-

alised linear models (GLM) with binomial error for proportion data (Crawley, 2012).

A different pipeline could be followed to describe diversity without any previous knowledge on DNA sequences on bdelloid rotifers, adopting an uninformed approach in the delimitation of species. Confirming the reliability of this approach would suggest that biodiversity analyses through DNA sequence data could be performed even in the absence of a reference database (Leese *et al.*, 2018). The use of DNA sequence data in the DNA taxonomy of understudied taxa is quite developed, with several methods that have already been applied to microscopic animals. Among these methods, we selected the Automated Barcode Gap Discovery, ABGD (Puillandre *et al.*, 2012), which is known to be reliable in COI of rotifers (Mills *et al.*, 2017) using the default settings of Pmin and Pmax on uncorrected genetic distances. Our dataset includes several individuals from few morphological species, limiting the problems in using ABGD with incompletely sampled taxa (Ahrens *et al.*, 2016). The ABGD approach identifies the best delineation of taxonomic units



**Figure 1.** Examples of voucher photographs of the identified individuals of bdelloid rotifers that were processed for DNA extraction. A: *Dissotrocha aculeata* D10\_DA01a; B: *Habrotrocha lata* D01\_HL01c; C: *Dissotrocha macrostyla* D01\_DM03a. Ejemplos de fotografías de especímenes de referencia de los individuos identificados de rotíferos bdelloideos que se usaron para extracción de ADN. A: *Dissotrocha aculeata* D10\_DA01a; B: *Habrotrocha lata* D01\_HL01c; C: *Dissotrocha macrostyla* D01\_DM03a.

potentially equivalent to species on the basis of the clearest barcoding gap between them and it is unlinked to the availability of previous taxonomic knowledge. We assessed whether the units of diversity discovered by ABGD matched the morphological species or not. We then asked whether the number of ABGD units for each morphological species could be due to the number of individuals or of populations for each species. We addressed this issue by using generalised linear models (GLM) with quasipoisson error for count data (Crawley, 2012).

As a description of the genetic variability in bdelloid rotifers from *Sphagnum* bogs, we provided metrics of uncorrected genetic distances within and between taxonomic units. All analyses were performed in R 3.3.3 (R Core Team, 2017), with the package ape v5.0 for handling DNA sequence data (Paradis *et al.*, 2004).

## RESULTS

In total, 17 morphological species of bdelloid rotifers were identified from the 13 samples (Fig. 1, Table 2), and 116 COI sequences were obtained (GenBank accession numbers MH251750-MH251865; Table S1, see Supplementary information, available at <http://www.limnetica.net/en/limnetica>). Out of the 109 animals identified to species level, 92 provided a correct species identification with BLAST searches against GenBank: thus, in 84.5 % of the cases, the highest match indeed identified the same name of the morphological species. The cases of mismatch referred to five morphological species, namely (see Supplementary information, Table S1).

For *Dissotrocha macrostyla*, most of the animals, 11 out of 14 (78.6 %), had the highest similarity to sequences of another species of the same genus, *D. aculeata* (see Supplementary information, Table S1). The genetic distance to the closest match was significantly smaller for the correctly identified animals (0.6 to 13.4 %) than for the incorrectly identified ones (13.2 to 14.3 %) (ANOVA:  $F_{1,12} = 6.0$ ,  $p = 0.03$ ).

For *Habrotrocha lata*, one animal out of three was not correctly assigned and had the closest match to a species of another family (*Pleuretra lineata*, family Philodinidae: see Supplementary

information, Table S1). The correctly identified sequences had the closest match at genetic distances of 6.5 and 7.5 %, whereas the incorrectly identified one was at higher distance, 11.3 %.

For *Macrotrachela plicata*, none of the two animals was correctly assigned. The closest match corresponded to other species of the same family (see Supplementary information, Table S1), with genetic distances of 10.2 and 10.7 %.

For *Otostephanos donneri*, the only animal was incorrectly assigned to a species of a different family (see Supplementary information, Table S1), with a distance of 12.4 %.

For *Rotaria macroceros*, none of the two animals was correctly assigned. One was assigned to *R. rotatoria*, a species of the same genus (see Supplementary information, Table S1), with a distance of 10.8 % and the other to a species of a different family, with a distance of 12.1 %.

All other species were correctly assigned (see Supplementary information, Table S1), even the ones with several animals from different populations (Table 2). Overall, the genetic distance to the closest match was significantly smaller for the correctly identified animals (average: 5.2 %, range: 0.002 to 15.3 %) than for the incorrectly identified ones (12.9 %, 10.2 to 14.3 %) ( $F_{1,114} = 44.3$ ,  $p < 0.0001$ ). The length of the overlapping part of the sequences with the ones in GenBank was not significantly different between correctly and incorrectly identified animals ( $F_{1,114} = 0.1$ ,  $p = 0.92$ ). The proportion of identified individuals for each species was not affected by any of the included variables: neither by the number of animals sequenced for each species (GLM:  $z = -0.1$ ,  $p = 0.90$ ), nor by the number of sequences available in GenBank ( $z = 0.0$ ,  $p = 0.13$ ).

Using a taxonomically blind approach without the use of a reference database, the 116 COI sequences provided evidence of 31 ABGD taxonomic units from the 17 morphological species (Table 2): the barcoding gap identified in the whole dataset by the application of ABGD was between 3.0 % (maximum intra-unit genetic distance) and 8.0 % (minimum inter-unit genetic distance) (Table 3). While a barcoding gap existed between the 31 ABGD taxonomic units, no clear barcoding gap was visible between the 17 morphological species. For them, the minimum

**Table 3.** Intraspecific minimum, maximum and mean uncorrected genetic distances between the 31 ABGD taxonomic units, with sample size (N). Intraspecific values are not reported for units with only 1 individual and NA means that no distance could be calculated between animals because they had the same COI sequence. The maximum genetic distance within each ABGD unit was 0.030, on average 0.011. The distances between all ABGD units were minimum = 0.080, maximum = 0.242, mean = 0.161. *Distancias genéticas sin corregir intraespecíficas mínimas, máximas y medianas entre las 31 unidades taxonómicas ABGD, con el número de especímenes N. Los valores intraespecíficos no se calcularon para unidades con un solo individuo. NA significa que la distancia no se pudo calcular ya que todos los animales tenían la misma secuencia de COI. La distancia genética máxima entre cada unidad ABGD era de 0.030, con un promedio de 0.011. Las distancias entre todas las unidades ABGD eran de mínimo 0.080, máximo 0.242 y una media de 0.161.*

ABGD unit	minimum	maximum	mean	N
s02	0.002	0.008	0.004	11
s03	NA	NA	NA	3
s05	0.005	0.028	0.015	11
s06	0.005	0.030	0.017	8
s07	0.004	0.004	0.004	3
s08	NA	NA	NA	2
s09	0.002	0.024	0.008	18
s10	0.002	0.010	0.007	3
s12	0.002	0.025	0.015	8
s13	0.006	0.006	0.006	2
s14	0.006	0.006	0.006	2
s16	0.003	0.025	0.012	6
s19	0.002	0.003	0.002	4
s20	0.003	0.003	0.003	3
s21	0.002	0.002	0.002	3
s23	NA	NA	NA	2
s26	0.022	0.022	0.022	2
s30	0.016	0.029	0.021	9
s31	0.003	0.012	0.007	4

inter-specific genetic distance was still 8.0 % but the maximum intra-specific genetic distance was much higher (Table 4): 10.4 % for *M. quadricornifera*, 11.5 % for *P. citrina*, 12.8 % for *H. lata*, 13.5 % for *Rotaria* sp., 13.8 % for *R. macroceros*, 15.6 % for *D. macrostyla*, and even 19.5 % for *R. rotatoria*. The intraspecific value of 19.5 % for *R. rotatoria* is only slightly smaller than the maximum difference in the whole dataset of 116 sequences for all the bdelloids, 24.2 %.

Most of the morphological species with several individuals were split into several ABGD taxonomic units: the highest number was eight for *R. rotatoria*, represented by 48 animals, and three for *D. macrostyla*, represented by 14 animals (Table 2). The number of ABGD taxonomic units for each morphological species was strongly biased by the number of sequences for each morphological species (GLM:  $z = 4.9$ ,  $p = 0.0002$ ) but not by the number of different populations for each morphological species ( $z = 0.3$ ,  $p = 0.77$ ). Indeed, even though different ABGD

units within the same morphological species were often found in different samples, these occurred also in a few cases in the same population. The most extreme case is that of the only two individuals of *R. macroceros* found in sample D10, which belonged to two different ABGD taxonomic units, s27 and s28 (Table 2) with a genetic distance of 13.8 % between them. The other instance of co-occurring ABGD taxonomic units within the same morphological species was for *R. rotatoria*, with two ABGD taxonomic units found in sample D03 (s09 and s11, 8.3-8.7 % distance between them), and even four ABGD units in sample D08 (s20, s21, s22, and s23, with 11.6-19.5 % distance between them).

## DISCUSSION

The main result of our DNA barcoding survey of bdelloid rotifers from *Sphagnum* bogs in the Swiss Jura Mountains is that the approach provides rather consistent estimates of diversity

**Table 4.** Intraspecific minimum, maximum and mean uncorrected genetic distances between the 17 morphological species, with sample size (N). Intraspecific values are not reported for species with only one individual. In addition, the number of ABGD taxonomic units is reported for each species. The maximum genetic distance within each species was 0.195, on average 0.109. The distances between species were minimum = 0.080, maximum = 0.242, mean = 0.164. *Distancias genéticas sin corregir mínimas, máximas y media entre las 17 especies morfológicas, con número de individuos (N). Los valores intraespecíficos no se registraron para especies con un solo individuo. Asimismo, se reportó el número de unidades taxonómicas ABGD para cada especie. La distancia genética máxima dentro de cada especie era de 0.195, con un promedio de 0.109. Las distancias entre todas las especies eran de mínimo 0.080, máximo 0.242 y una media de 0.164.*

species	minimum	maximum	mean	N	ABGD
<i>Adineta gracilis</i>	0.002	0.025	0.015	8	1
<i>Dissotrocha aculeata</i>	0.022	0.022	0.022	2	1
<i>Dissotrocha macrostyla</i>	0.002	0.156	0.068	14	3
<i>Habrotrocha lata</i>	0.006	0.128	0.082	3	2
<i>Macrotrachela plicata</i>	0.006	0.006	0.006	2	1
<i>Macrotrachela quadricornifera</i>	0.003	0.105	0.068	7	2
<i>Philodina citrina</i>	0.003	0.115	0.038	7	2
<i>Rotaria macroceros</i>	0.138	0.138	0.138	2	2
<i>Rotaria magnacalcarata</i>	0.002	0.003	0.002	4	1
<i>Rotaria rotatoria</i>	0.002	0.195	0.120	48	8
<i>Rotaria</i> sp.	0.004	0.135	0.087	6	2
<i>Rotaria tardigrada</i>	0.005	0.030	0.017	8	1

with the traditional morphological methods. Overall, most of the DNA sequences of COI from the sequenced animals matched to the correct species name in the GenBank reference database. The 17 sequences (15.6 %) that did not match the correct species names corresponded to sequences that did not have any close match in GenBank and were thus incorrectly assigned: all of these incorrectly assigned animals had a genetic distance to the closest match above 10 %, which is a rather high genetic distance in COI for animals, even for rotifers (Tang *et al.*, 2014). A COI distance of 10 % is indeed typically above the barcoding gap used to separate two closely related species in many groups of animals (Hebert *et al.*, 2003).

In rotifers, it is known that animals of the same morphospecies may exceed 10 % in their genetic distances in COI (Fontaneto, 2014). Such high genetic distances within the same morphospecies is known in rotifers for cases of cryptic or pseudocryptic species (e.g. Mills *et al.*, 2017; Moreno *et al.*, 2017). Indeed, the use of a statistical approach to identify taxonomic units

from DNA sequence data, such as the ABGD method we applied, revealed that several cryptic taxa could be potentially present in our dataset. Out of 17 morphological species, 31 ABGD units were found, with minimum genetic distances between them of 8 %. Such a threshold is lower than the 10 % distance of the incorrectly assigned sequences. Thus, we can support the hypothesis that all the misidentification we had were due to the occurrence of cryptic species coupled with the lack of corresponding DNA sequence information in GenBank for each cryptic species within the complexes.

We are confident that further studies would be able to fill this knowledge gap in the reference database: one of the extreme cases of the occurrence of cryptic species in bdelloid rotifers is *Rotaria rotatoria*, with an estimated number of few tens of species in the complex (Fontaneto *et al.*, 2009). Yet, because of a good representativeness of the species complex in the reference database, with more than 800 sequences already available in GenBank, none of the 48 animals of this

morphospecies in our dataset was incorrectly identified. More taxonomic work should be performed for the morphospecies *R. rotatoria* in order to resolve the taxonomic ambiguity, following what was done for the emblematic case of the *Brachionus plicatilis* species complex (Mills *et al.*, 2017) for which 15 species were determined from integrative taxonomic approaches combining extensive barcoding, morphology, and also geographic information. Yet, notwithstanding the high taxonomic uncertainties for *R. rotatoria*, the survey we performed based on DNA sequence assigned all animals to the correct species complex, and the new genetic information provided by our survey further increased the reference database.

The lack of corresponding sequences in GenBank can be filled only by further faunistic and taxonomic studies including DNA sequence information, similar to the one we presented here. The need for a reliable reference database is one of the optimal requirements for biological monitoring of aquatic habitats through DNA barcoding and metabarcoding (Leese *et al.*, 2018). The endeavour of obtaining and managing such a database started several years ago with shared information through GenBank and BOLD, and for some groups it developed in taxonomically curated and reliable systems to query the sequences obtained from the field, for example in prokaryotes (SILVA, Quast *et al.*, 2013), in protists (UniEuk, Berney *et al.*, 2017), and in fungi (UNITE, Abarenkov *et al.*, 2010). Any metabarcoding study on protists sequenced from the field in bulk extractions of organisms or from environmental DNA has the very useful UniEuk system as a reference for protist species, but the same study will recover sequences from rotifers and from other microscopic animals such as nematodes, tardigrades, and gastrotrichs, which are of similar size as several protists and live in the same habitats. Thus, it would be useful to start a curated reference system also for microscopic animals, or even for them to be included in the UniEuk reference database for unicellular eukaryotes.

Overall, we can conclude that we are ready to assign bdelloid species identification starting from DNA sequence information in aquatic habitats. We were successful in spite of the fact that we focused on a previously understudied country

and habitat for rotifers. Our suggestion for future applications is to adopt a confident approach and trust only close matches that are lower than 10 % in genetic distances, while leaving as unidentified all sequences that have a higher genetic distance. We are confident that in the future the representativeness of GenBank, BOLD, or any other dedicated reference system will improve, but at least for acidic aquatic habitats of Central Europe we demonstrated that the approach could be considered doable and reliable already now.

## ACKNOWLEDGEMENTS

We thank the relevant authorities for granting us access to the protected peatlands for sampling, and the Swiss Federal Office for the Environment (FOEN) for funding this work through the Swiss-BOL - Swiss Inventory of Genetic Biodiversity project.

## REFERENCES

- ABARENKOV, K., R. HENRIK NILSSON, K. H. LARSSON, I. J. ALEXANDER, U. EBERHARDT, S. ERLAND, K. HØILAND, R. KJØLLER, E. LARSSON, T. PENNANEN, R. SEN *et al.* 2010. The UNITE database for molecular identification of fungi—recent updates and future perspectives. *New Phytologist*, 186(2): 281-285. DOI: 10.1111/j.1469-8137.2009.03160.x
- AHRENS, D., T. FUJISAWA, H.-J. KRAMMER, J. EBERLE, S. FABRIZI, A.P. VOGLER. 2016. Rarity and incomplete sampling in DNA-based species delimitation. *Systematic Biology*, 65 (3): 478-494. DOI: 10.1093/sysbio/syw002
- BAIRD, D. J. & M. HAJIBABAEI. 2012. Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology*, 21(8): 2039-2044. DOI: 10.1111/j.1365-294X.2012.05519.x
- BENSON, D. A., M. CAVANAUGH, K. CLARK, I. KARSCH-MIZRACHI, D. J. LIPMAN, J. OSTELL & E. W. SAYERS. 2013. GenBank. *Nucleic Acids Research*, 41(D1): D36–D42. DOI: 10.1093/nar/gkt1030

- BERNEY, C., A. CIUPRINA, S. BENDER, J. BRODIE, V. EDGCOMB, E. KIM, J. RAJAN, L. WEGENER PARFREY, S. ADL, S. AUDIC, D. BASS, *et al.* 2017. UniEuk: time to speak a common language in protistology!. *Journal of Eukaryotic Microbiology*, 64(3): 407-411. DOI: 10.1111/jeu.12414
- CAMPILLO, S., E. M. GARCIA-ROGER, D. MARTINEZ-TORRES & M. SERRA. 2005. Morphological stasis of two species belonging to the L-morphotype in the *Brachionus plicatilis* species complex. *Hydrobiologia*, 546: 181-187. DOI: 10.1007/s10750-005-4120-7
- CRAWLEY, M.J. 2012. *The R book*. 2<sup>nd</sup> edition. John Wiley & Sons.
- DE JONG, Y., M. VERBEEK, V. MICHELSEN, P. DE PLACE BJØRN, W. LOS, F. STEEMAN, N. BAILLY, C. BASIRE, P. CHYLARECKI, E. STLOUKAL, G. HAGEDORN *et al.* 2014 Fauna Europaea - all European animal species on the web. *Biodiversity Data Journal*, 2: e4034. DOI: 10.3897/BDJ.2.e4034
- DÍAZ, J., G. V. VILLANOVA, F. BRANCOLINI, F. DEL PAZO, V. M. POSNER, A. GRIMBERG & S. E. ARRANZ. 2016. First DNA barcode reference library for the identification of South American freshwater fish from the lower Paraná river. *PLoS ONE*, 11(7): e0157419. DOI: 10.1371/journal.pone.0157419
- DONNER, J. 1965. *Ordnung Bdelloidea (Rotatoria, Radertiere)*. Akademie Verlag.
- DUMONT, H. & H. SEGERS. 1996. Estimating lacustrine zooplankton species richness and complementarity. *Hydrobiologia*, 341 (1): 125-132. DOI: 10.1007/BF00018116
- FOLMER, O., M. BLACK, W. HOEH, R. LUTZ & R. VRIJENHOEK. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3: 294-299.
- FONTANETO, D. 2014. Molecular phylogenies as a tool to understand diversity in rotifers. *International Review of Hydrobiology*, 99(1-2): 178-187. DOI: 10.1002/iroh.201301719
- FONTANETO, D., M. KAYA, E. A. HERNIOU & T.G. BARRACLOUGH. 2009. Extreme levels of hidden diversity in microscopic animals (Rotifera) revealed by DNA taxonomy. *Molecular Phylogenetics and Evolution*, 53(1): 182-189. DOI: 10.1016/j.ympev.2009.04.011
- FONTANETO, D., A. M. BARBOSA, H. SEGERS & M. PAUTASSO. 2012. The 'rotiferologist' effect and other global correlates of species richness in monogonont rotifers. *Ecography*, 35(2): 174-182. DOI: 10.1111/j.1600-0587.2011.06850.x
- FONTANETO, D. & W. DE SMET. 2015. *Rotifera*. In: Handbook of zoology, Gastrotricha and Gnathifera (ed A Schmidt-Rhaesa), pp. 216-300.
- GARCÍA-MORALES, A. E. & M. ELÍAS-GUTIÉRREZ. 2013. DNA barcoding of freshwater Rotifera in Mexico: Evidence of cryptic speciation in common rotifers. *Molecular Ecology Resources*, 13(6): 1097-1107. DOI: 10.1111/1755-0998.12080
- GILBERT, J. J. 2017. Non-genetic polymorphisms in rotifers: environmental and endogenous controls, development, and features for predictable or unpredictable environments. *Biological Reviews*, 92(2): 964-992. DOI: 10.1111/brv.12264
- GÓMEZ, A., M. SERRA, G. R. CARVALHO & D. H. LUNT. 2002. Speciation in ancient cryptic species complexes: evidence from the molecular phylogeny of *Brachionus plicatilis* (Rotifera). *Evolution*, 56(7): 1431-1444. DOI: 10.1111/j.0014-3820.2002.tb01455.x
- HEBERT, P.D., A. CYWINSKA & S. L. BALL. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1512): 313-321. DOI: 10.1098/rspb.2002.2218
- HUEMER, P., M. MUTANEN, K. M. SEFC & P. D. HEBERT. 2014. Testing DNA barcode performance in 1000 species of European Lepidoptera: large geographic distances have small genetic impacts. *PLoS ONE*, 9(12): e115774. DOI: 10.1371/journal.pone.0115774
- KATOH, K. & D. M. STANDLEY. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4): 772-780. DOI: 10.1093/molbev/mst010

- KORDBACHEH, A., G. GARBALENA & E. J. WALSH. 2017. Population structure and cryptic species in the cosmopolitan rotifer *Euchlanis dilatata*. *Zoological Journal of the Linnean Society*, 181(4): 757-777. DOI: 10.1093/zoolinnean/zlx027
- KUCZYŃSKA-KIPPEN, N. 2018. The use of bdelloids in reference to rotifer biocoenotic indices as an indicator of the ecological state of small field water bodies: The effect of macrophytes, shading and trophic state of water. *Ecological Indicators*, 89: 576-583. DOI: 10.1016/j.ecolind.2018.02.046
- LEESE, F., A. BOUCHEZ, K. ABARENKOV, F. ALTERMATT, Á. BORJA, K. BRUCE, T. EKREM, F. ČIAMPOR, Z. ČIAMPOROVÁ-ZAŤOVIČOVÁ, F. O. COSTA, S. DUARTE, *et al.* 2018. Why we need sustainable networks bridging countries, disciplines, cultures and generations for aquatic biomonitoring 2.0: a perspective derived from the DNAqua-net COST Action. *Advances in Ecological Research*, 58: 63-99. DOI: 10.1016/bs.aecr.2018.01.001
- MADDISON, W. P. & D. R. MADDISON. 2018. *Mesquite: a modular system for evolutionary analysis*. Version 3.40 <http://mesquiteproject.org>
- MILLS, S., J. A. ALCÁNTARA-RODRÍGUEZ, J. CIROS-PÉREZ, A. GÓMEZ, A. HAGIWARA, K. H. GALINDO, C. D. JERSABEK, R. MALEKZADEH-VIAYEH, F. LEASI, J. S. LEE, D. B. MARK WELCH, *et al.* 2017. Fifteen species in one: deciphering the *Brachionus plicatilis* species complex (Rotifera, Monogononta) through DNA taxonomy. *Hydrobiologia*, 796(1): 39-58. DOI: 10.1007/s10750-016-2725-7
- MORENO, E., J. M. CONDE-PORCUNA & A. GÓMEZ. 2017. Barcoding rotifer biodiversity in Mediterranean ponds using diapausing egg banks. *Ecology and Evolution*, 7(13): 4855-4867. DOI: 10.1002/ece3.2986
- MORINIÈRE, J., L. HENDRICH, M. BALKE, A. J. BEERMANN, T. KÖNIG, M. HESS, S. KOCH, R. MÜLLER, F. LEESE, P. D. HEBERT & A. HAUSMANN. 2017. A DNA barcode library for Germany's mayflies, stoneflies and caddisflies (Ephemeroptera, Plecoptera and Trichoptera). *Molecular Ecology Resources*, 27: 1755-0998. DOI: 10.1111/1755-0998.12683
- OBERTEGGER, U., H. A. SMITH, G. FLAIM & R. L. WALLACE. 2011. Using the guild ratio to characterize pelagic rotifer communities. *Hydrobiologia*, 662(1): 157-162. DOI: 10.1007/s10750-010-0491-5
- OBERTEGGER, U., A. CIEPLINSKI, D. FONTANETO & S. PAPAKOSTAS. 2018. Mitonuclear discordance as a confounding factor in the DNA taxonomy of monogonont rotifers. *Zoologica Scripta*, 47(1): 122-132. DOI: 10.1111/zsc.12264
- PAPAKOSTAS, S., E. MICHALOUDI, K. PROIOS, M. BREHM, L. VERHAGE, J. ROTA, C. PEÑA, G. STAMOU, V. L. PRITCHARD, D. FONTANETO & S. A. DECLERCK. 2016. Integrative taxonomy recognizes evolutionary units despite widespread mitonuclear discordance: evidence from a rotifer cryptic species complex. *Systematic Biology*, 65(3): 508-524. DOI: 10.1093/sysbio/syw016
- PARADIS E., J. CLAUDE & K. STRIMMER. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20: 289-290. DOI: 10.1093/bioinformatics/btg412
- PULLANDRE, N., A. LAMBERT, S. BROUILLET & G. ACHAZ. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, 21(8): 1864-1877. DOI: 10.1111/j.1365-294X.2011.05239.x
- QUAST, C., E. PRUESSE, P. YILMAZ, J. GERKEN, T. SCHWEER, P. YARZA, J. PEPLIES & F. O. GLÖCKNER. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41 (D1): D590-D596. DOI: 10.1093/nar/gks1219
- R CORE TEAM. 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org>.
- SEGRS, H. 2007. Annotated checklist of the rotifers (Phylum Rotifera), with notes on nomenclature, taxonomy and distribution. *Zootaxa*, 1564: 1-104. DOI: 10.11646/zootaxa.1564.1.1

- SEGERS, H. & W. H. DE SMET. 2008. Diversity and endemism in Rotifera: a review, and Keratella Bory de St Vincent. *Biodiversity Conservation*, 17: 303–316. DOI: 10.1007/s10531-007-9262-7
- SLÁDEČEK, V. 1983. Rotifers as indicators of water quality. *Hydrobiologia*, 100: 169-201. DOI: 10.1007/BF00027429
- SUATONI, E., S. VICARIO, S. RICE, T. SNELL & A. CACCONE, 2006. An analysis of species boundaries and biogeographic patterns in a cryptic species complex: the rotifer—*Brachionus plicatilis*. *Molecular Phylogenetics and Evolution*, 41(1): 86-98. DOI: 10.1016/j.ympev.2006.04.025
- TABERLET, P., E. COISSAC, F. POMPANON, C. BROCHMANN & E. WILLERSLEV. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, 21(8): 2045-2050. DOI: 10.1111/j.1365-294X.2012.05470.x
- TANG, C.Q., U. OBERTEGGER, D. FONTANETTO & T. G. BARRACLOUGH. 2014. Sexual species are separated by larger genetic gaps than asexual species in rotifers. *Evolution*, 68(10): 2901-2916. DOI: 10.1111/evo.12483.