THE ROYAL SOCIETY biology letters

DNA taxonomy of a neglected animal phylum: an unexpected diversity of tardigrades

Mark Blaxter*, Ben Elsworth and Jennifer Daub

Institute of Cell, Animal and Population Biology, University of Edinburgh, King's Buildings, Edinburgh EH9 3JT, UK
**Author for correspondence (mark.blaxter@ed.ac.uk).*

Recd 09.10.03; *Accptd* 29.10.03; *Published online* 19.12.03

A molecular survey technique was used to investigate the diversity of terrestrial tardigrades from three sites within Scotland. Ribosomal small subunit sequence was used to classify specimens into molecular operational taxonomic units (MOTU). Most MOTU were identified to the generic level using digital voucher photography. Thirty-two MOTU were defined, a surprising abundance given that the documented British fauna is 68 species. Some tardigrade MOTU were shared between the two rural collection sites, but no MOTU were found in both urban and rural sites, which conflicts with models of ubiquity of meiofaunal taxa. The patterns of relatedness of MOTU were particularly intriguing, with some forming clades with low levels of divergence, suggestive of taxon flocks. Some morphological taxa contained well-separated MOTU, perhaps indicating the existence of cryptic taxa. DNA sequence-based MOTU proved to be a revealing method for meiofaunal diversity studies.

Keywords: DNA; barcoding; molecular; taxonomy; tardigrade; biodiversity

1. INTRODUCTION

Estimates of overall taxon diversity of animal life on Earth are in the tens of millions, but only *ca*. 1.5 million animal species have been formally described (May 1988; de Meeus & Renaud 2002). This taxonomy deficit is especially marked for microfauna and meiofauna. These taxa are problematic to diagnose and often hyperabundant, making identification of rare novelty a Herculean task. A universal DNA barcode system for animal species diagnosis could provide detailed and testable hypotheses of specimen membership of taxa, and of relationships of taxa to each other. These would be portable between both taxon groups and individual researchers, and adaptable to high throughput (Floyd *et al.* 2002; Hebert *et al.* 2003*b*). Two genes have been proposed for DNA barcoding: the nuclear small subunit ribosomal RNA gene (SSU; Floyd *et al.* 2002) and the mitochondrial cytochrome *c* oxidase subunit 1 gene (COX1; Hebert *et al.* 2003*b*). Both have a high copy number per cell and conserved primer-binding sites for easy amplification, and patterns of variability that include slowly and rapidly evolving segments. SSU sequence data can distinguish morphologically cryptic taxa of nematodes (Eyualem & Blaxter 2003) and have been employed in random surveys (Floyd *et al.* 2002). The COX1 gene has been used to survey hexapods (Hebert *et al.* 2003*b*) and in theoretical analyses of animal diversity (Hebert *et al.* 2003*a*).

Tardigrades (phylum Tardigrada) are widely distributed in marine and terrestrial sediments, and in moss and lichen (hence the common name moss-bears; Kinchin 1994; Goldstein & Blaxter 2002). They range in size from less than 0.1 mm to *ca*. 1.2 mm. Tardigrada is not speciose (*ca*. 900 described taxa; Morgan & King 1976; Maucci 1986; Kinchin 1994). In the British Isles, the tardigrade fauna has been studied by Murray (1905, 1906, 1911) and Morgan & King (1976): 68 terrestrial species have been identified (Morgan & King 1976). Distribution maps suggest that surveys have not been exhaustive, with most of the UK devoid of records (Morgan & King 1976); Murray in particular sampled many Scottish sites (Murray 1905, 1906, 1911). We have applied a DNA barcode system (Floyd *et al.* 2002) to tardigrades from urban and rural sites in southern Scotland and defined molecular operational taxonomic units (MOTU) that we can correlate with morphological species through digital image vouchers. We find a surprising diversity of tardigrade taxa, including cryptic morphological taxa.

2. MATERIAL AND METHODS

Tardigrades were collected from 12 locations within Edinburgh city from moss cushions on rock and soil substrates. Mosses and lichens from standing trees and drystone walls in two rural sites were also surveyed: the Wood of Cree, Dumfries and Galloway—an ancient native oak woodland recovering from clear felling in *ca*. 1920 (five samples)—and Ettrick old churchyard, Glen Ettrick (nine samples). Over 300 specimens were isolated. Sequencing was focused on specimens showing morphological disparity within each sample. Tardigrades were collected into 5 µl of sterilized tap water in a vinyl microplate and a digital image voucher was collected. Each specimen was digested in 20 µl of 0.25 M NaOH (Floyd *et al.* 2002). After neutralization, 1–2 µl of the extract were taken to a PCR reaction with primers SSU_18A and SSU_26R (Blaxter *et al.* 1998). PCR products (*ca*. 1 kb) were cleaned with shrimp alkaline phosphatase and exonuclease I, and sequenced with primer SSU_9R and ABI BigDye 3.1 reagents on an ABI 3730 sequencer. The sequences have been deposited in EMBL/GenBank under the accession numbers AJ617376–AJ617468.

Sequence traces were basecalled using PHRED (Ewing *et al.* 1998), screened for size (more than 475 bp) and clustered into MOTUs using a perl script, define_motu.pl (outlined in figure 3 in electronic Appendix A, available on The Royal Society's Publications Web site; R. Floyd, unpublished data), based on CLOBB (Parkinson *et al.* 2002). This program clusters sequences on the basis of identity using a specified cut-off. We used a 2 bp cut-off (i.e. 0.4% difference over 500 bases) based on determined experimental error (Floyd *et al.* 2002). Unresolved bases were ignored for purposes of MOTU classification. The sequences were aligned to tardigrade SSU sequences from EMBL/GenBank (Garey *et al.* 1996; Moon & Kim 1996) and from a survey of cultured tardigrades. Three sequences from unidentified Antarctic soil meiofauna were supplied by A. Rogers, L. Peck and P. Convey, British Antarctic Survey, Cambridge. The alignment was analysed using PAUP v. 4.b10 (Swofford 1999) and MRBAYES (Huelsenbeck & Ronquist 2001). For Bayesian analysis, trees were sampled every 100 generations of a million-generation Markov chain run, and the last 9000 trees used for calculation of posterior probabilities. A relational database of all specimen metadata was used for analyses of data correlations.

3. RESULTS AND DISCUSSION

Ninety-five SSU sequences were generated from 116 positive PCR reactions. The sequences clustered into 32 MOTU. None was identical to the seven tardigrade SSUs from EMBL/GenBank (figure 1*a*). Sequences from the Antarctic and cultured tardigrade isolates were also unique in the dataset. Phylogenetic analysis of representative sequences from each MOTU yielded a well-resolved

DOI 10.1098/rsbl.2003.0130

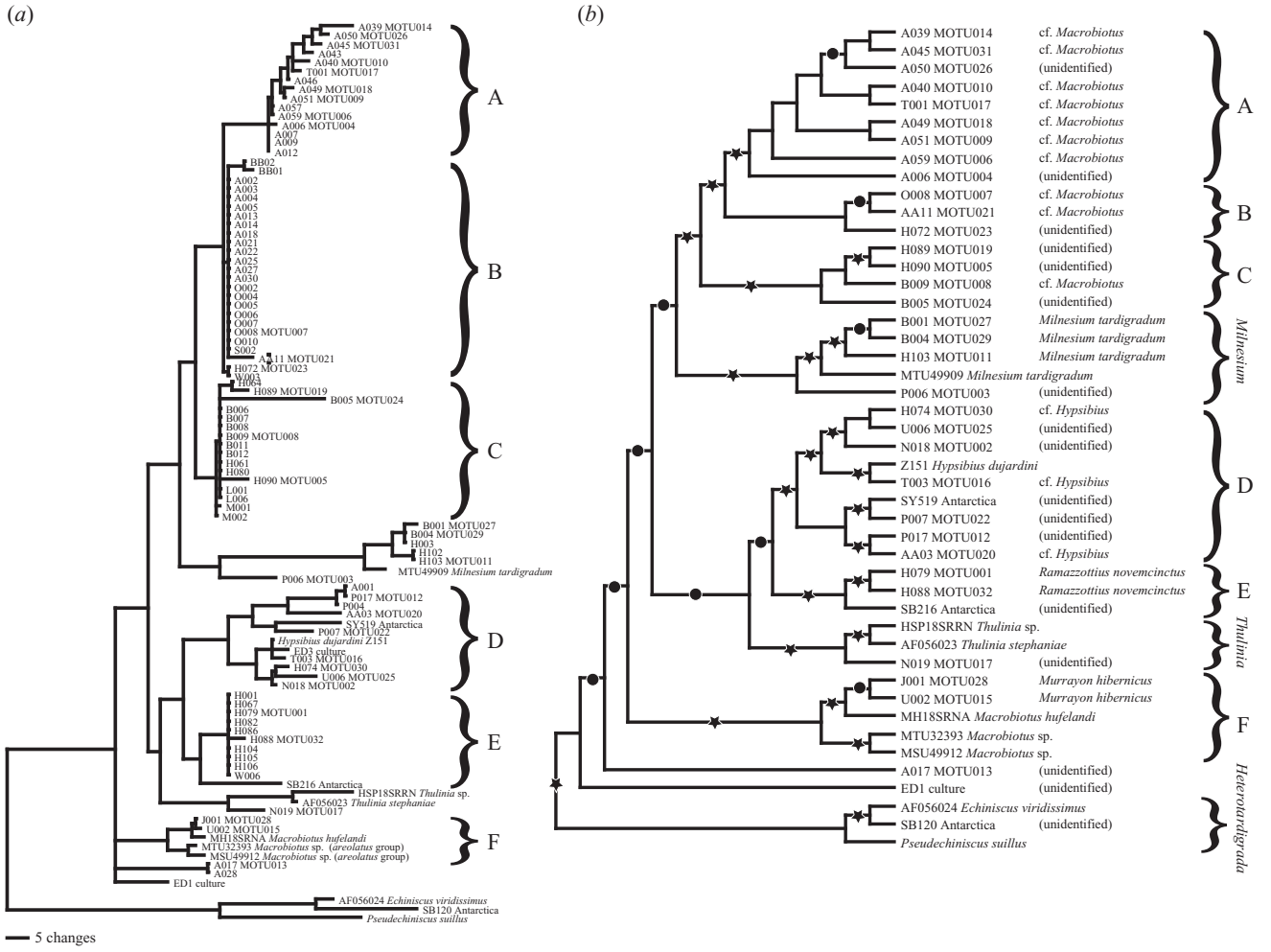

Figure 1. Phylogenetic analysis of tardigrade SSU data. (*a*) The DNA barcode sequences (all greater than 475 bp) were aligned and analysed using the neighbour joining algorithm (Swofford 1999), with absolute difference as a pairwise measure. The alignment length used was 510 bp. The clustering of individual sequences in the tree indicates the membership of each MOTU (but note that MOTU were inferred from a tree-independent clustering process; see figure 3 in electronic Appendix A). (*b*) An alignment of the longest sequences from each MOTU and other tardigrade sequences was analysed using Bayesian inference (Huelsenbeck & Ronquist 2001). Posterior probabilities associated with each node are indicated. The putative morphological identifications of specimens in each MOTU are given in italics to the right. Node support: stars indicate a greater than 99% posterior probability and circles denote a greater than 90% posterior probability.

tree with most nodes having significant posterior probabilities (figure 1*b*). MOTU were associated with known taxa by two methods. First, their phylogenetic relatedness to database sequences was assessed. We found, as has been described earlier (Garey *et al.* 1996), that the sequence HSP18SRRN, attributed to *Hypsibius* sp., robustly associates with AF056023 from *Thulinia stephaniae*. MOTU017 also associated with this clade, and thus was tentatively assigned to *Thulinia*. Similarly, MOTU011, MOTU027 and MOTU029 were closely related to *Milnesium tardigradum* MTU49909. A second method used digital vouchers to identify specimens. Most MOTU were assigned using morphological data, including several that have no previously sequenced close relatives. From these comparisons we are able to tentatively assign each MOTU to genus (figure 1*b*).

Several noteworthy patterns emerged. Clade B contains MOTU007 with 19 members with minimal within-MOTU variation, and a closely related taxon MOTU021. Clade D, including the cultured species *Hypsibius dujardini*, contained 10 well-separated MOTU. By contrast, clade A, was a comb-like series of nine closely related sister taxa with over 25 bases separating the most divergent members. These specimens were not simply diagnosed by a 2 bp rule as reclustering after randomizing addition order yielded between seven and nine MOTU (see electronic Appendix A, figure 3*b*). From manual inspection of the relevant traces the differences that separate them were not owing to sequencing error. Clade A included many specimens from site BE_A (a moss cushion from a west-facing stone wall in Ettrick), as did clade B MOTU007, suggesting that the pattern of sequence difference was owing to taxon biology rather than experimental error. If the cloud of MOTU in clade A is real, and there were indeed many independent lineages of closely related animals cohabiting one moss cushion, this challenges models of niche exclusion of congeneric taxa. Alternatively, the divergent SSU could instead be a novel biological feature of a single, very variable taxon. For example, these specimens could be from obligately asexual lineages that have accumulated differences by drift without sex. Indeed, many tardigrade taxa are asexual, with parthenogenesis

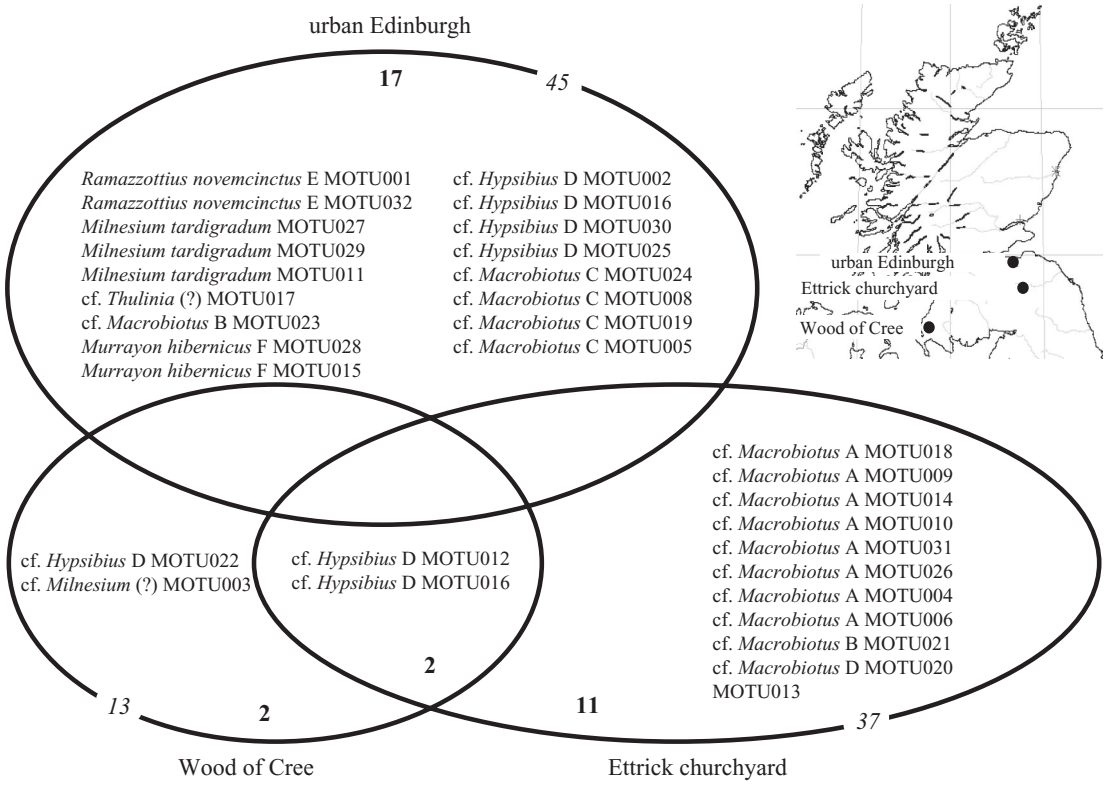

Figure 2. A disjoint distribution of tardigrade MOTU between the three sample sites. This Venn diagram shows the distribution of MOTU across the sample sites. The numbers of sequences per site are given in italic on each circle, and the MOTU in each segment are indicated in bold type. The inset indicates the geographical locations of the three sites in Scotland.

common in some genera (Kinchin 1994). Features such as this could render DNA barcoding approaches informatically challenging: heuristics for taxon delimitation might need to be conditional depending on known biology. *Milnesium tardigradum* is a distinctive tardigrade and to our knowledge only one species from the genus has been described from the UK to date. However, we found three MOTU with sequences close to the database *M. tardigradum* SSU, and with morphological identification as cf. *M. tardigradum*. As *M. tardigradum* is a sexual species, we suggest that these MOTU represent cryptic taxa, morphologically indistinguishable from *M. tardigradum sensu stricto*. Re-collection and closer examination is warranted.

One described feature of tardigrade faunas worldwide is the existence of globally distributed species. Tardigrades can undergo anhydrobiosis, and it has been proposed that individual species could be spread by the wind or another passive transport mechanism as dried tuns. We did not observe this pattern. Mapping MOTU onto sites (figure 2) revealed a disjunct distribution, with no MOTU shared between urban and rural sites. Out of 15 rural MOTU, only two were found in both sites. The cloud of taxa in clade A all derived from Ettrick, and all clade C specimens came from Edinburgh. Finlay has proposed that eukaryotic microbes lack biogeography, and that there exists a small number of ubiquitous taxa, a suggestion amply supported by data from protozoal diversity analyses (Finlay 2002). The proposed organismal size cut-off for ubiquity versus biogeography is *ca*. 1 mm. Most tardigrades fall below this cut-off, but our data from just three closely spaced sites suggest that tardigrades do have a biogeography. Global patterns of diversity might differ for multicellular, mainly sexual, taxa versus unicellular ones. DNA barcode sequencing of isolates from additional globally distributed species will determine whether they are indeed single taxa.

Overall, 32 MOTU were identified from only 95 sequences. Most MOTU included single specimens, and the collectors curve was still in its initial steep rise. Denser sampling is required to provide a reliable estimate of the diversity of true 2 bp MOTU. It seems unlikely that we have identified 40% of UK diversity (Morgan & King 1976) in so few samples, and we thus suggest that the UK tardigrade fauna is probably significantly underestimated.

The use of DNA barcodes for animal taxon identification has been vigorously debated (Blaxter 2003; Seberg *et al.* 2003; Tautz *et al.* 2003). Missing from much of this debate is a consideration and examination of sequence variability within traditionally defined taxa. Simply classifying each unique sequence as a new taxon is to ignore biological reality, but the amount of sequence difference that does indicate distinct taxon status for a group must be carefully defined (Floyd *et al.* 2002). The data presented here suggest that significant background data will be necessary for each taxon group studied to carefully define these metrics. However, in neglected and difficult major taxa such as the tardigrades it may be that DNA barcode data may be the only universal tool available, and that molecular operational taxonomic units will provide a useful starting point for further biological analyses.

Acknowledgements

This work was funded by the BBSRC and the Linnaean Society of London through the A. G. Side Fund. The authors thank A. Rogers,

L. Peck, P. Convey, R. Floyd and J. Lovell for data and support, and A. Davison and S. Nee for their comments on the manuscript.

Blaxter, M. 2003 Molecular systematics: counting angels with DNA. *Nature* **421**, 122–124.

Blaxter, M. L. (and 11 others) 1998 A molecular evolutionary framework for the phylum Nematoda. *Nature* **392**, 71–75.

de Meeus, T. & Renaud, F. 2002 Parasites within the new phylogeny of eukaryotes. *Trends Parasitol.* **18**, 247–251.

Ewing, B., Hillier, L., Wendl, M. C. & Green, P. 1998 Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**, 175–185.

Eyualem, A. & Blaxter, M. 2003 Comparison of biological, molecular and morphological methods of species identification in a set of cultured *Panagrolaimus* isolates. *J. Nematol.* **35**, 119–128.

Finlay, B. J. 2002 Global dispersal of free-living microbial eukaryote species. *Science* **296**, 1061–1063.

Floyd, R., Eyualem, A., Papert, A. & Blaxter, M. 2002 Molecular barcodes for soil nematode identification. *Mol. Ecol.* **11**, 839–850.

Garey, J. R., Krotec, M., Nelson, D. R. & Brooks, J. 1996 Molecular analysis supports a tardigrade–arthropod association. *Invert. Biol.* **115**, 79–88.

Goldstein, B. & Blaxter, M. 2002 Tardigrades. *Curr. Biol.* **12**, R475.

Hebert, P. D., Ratnasingham, S. & deWaard, J. R. 2003*a* Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond.* B **270**(Suppl.), S96–S99. (DOI 10.1098/rsbl.2003.0022.)

Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. 2003*b* Biological identifications through DNA barcodes. *Proc. R. Soc. Lond.* B **270**, 313–321. (DOI 10.1098/rspb.2002.2218.)

Huelsenbeck, J. P. & Ronquist, F. 2001 MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.

Kinchin, I. M. 1994 *The biology of tardigrades*. London: Portland Press.

Maucci, W. (ed.) 1986 *Fauna d'Italia Tardigrada*. Bologna: Edizioni Calderini Bologna.

May, R. M. 1988 How many species are there on earth? *Science* **241**, 1441–1449.

Moon, S. Y. & Kim, W. 1996 Phylogenetic position of the Tardigrada based on 18S ribosomal RNA gene sequences. *Zool. J. Linn. Soc.* **116**, 61–69.

Morgan, C. I. & King, P. E. 1976 *British tardigrades (Tardigrada): keys and notes for identification of the species*. London: Academic.

Murray, J. 1905 The Tardigrada of the Forth Valley. *Ann. Scottish Nat. Hist.* **55**, 160–164.

Murray, J. 1906 The Tardigrada of the Forth Valley. Part II. *Ann. Scottish Nat. Hist.* **60**, 214–217.

Murray, J. 1911 Scottish Tardigrada. A review of our present knowledge. *Ann. Scottish Nat. Hist.* **78**, 88–95.

Parkinson, J., Guiliano, D. & Blaxter, M. 2002 Making sense of EST sequences by CLOBBing them. *BMC Bioinf.* **3**, 31.

Seberg, O., Humphries, C. J., Knapp, S., Stevenson, D. W., Petersen, G., Scharff, N. & Andersen, N. M. 2003 Shortcuts in systematics? A commentary on DNA-based taxonomy *Trends Ecol. Evol.* **18**, 63–65.

Swofford, D. 1999 *PAUP* 4.b10*. Sunderland, MA: Sinauer.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. 2003 A plea for DNA taxonomy. *Trends Ecol. Evol.* **18**, 70–74.

Visit www.journals.royalsoc.ac.uk and navigate to this article through *Biology Letters* to see the accompanying electronic appendix.