Supplement S1 (Figs S1, S2, Table S1) to: Natta G. et al. 2024: DNA metabarcoding of gut microbiota reveals considerable taxonomic differences among wild individuals of the dung beetle Trypocopris pyrenaeus (Coleoptera: Geotrupidae). — *Eur. J. Entomol.* **121**: 40–53.

Collection of adult individuals of *T. pyrenaeus*

Adults of *T. pyrenaeus* were collected from four localities in the Western Alps, Piedmont, Italy, namely Bocchetto Sessera (BI), Colle del Lys (TO), Colle della Vaccera (TO) and Rorà (TO) and preserved in absolute ethanol for the molecular analyses. The individuals from the three last localities were only used for the barcoding analysis.

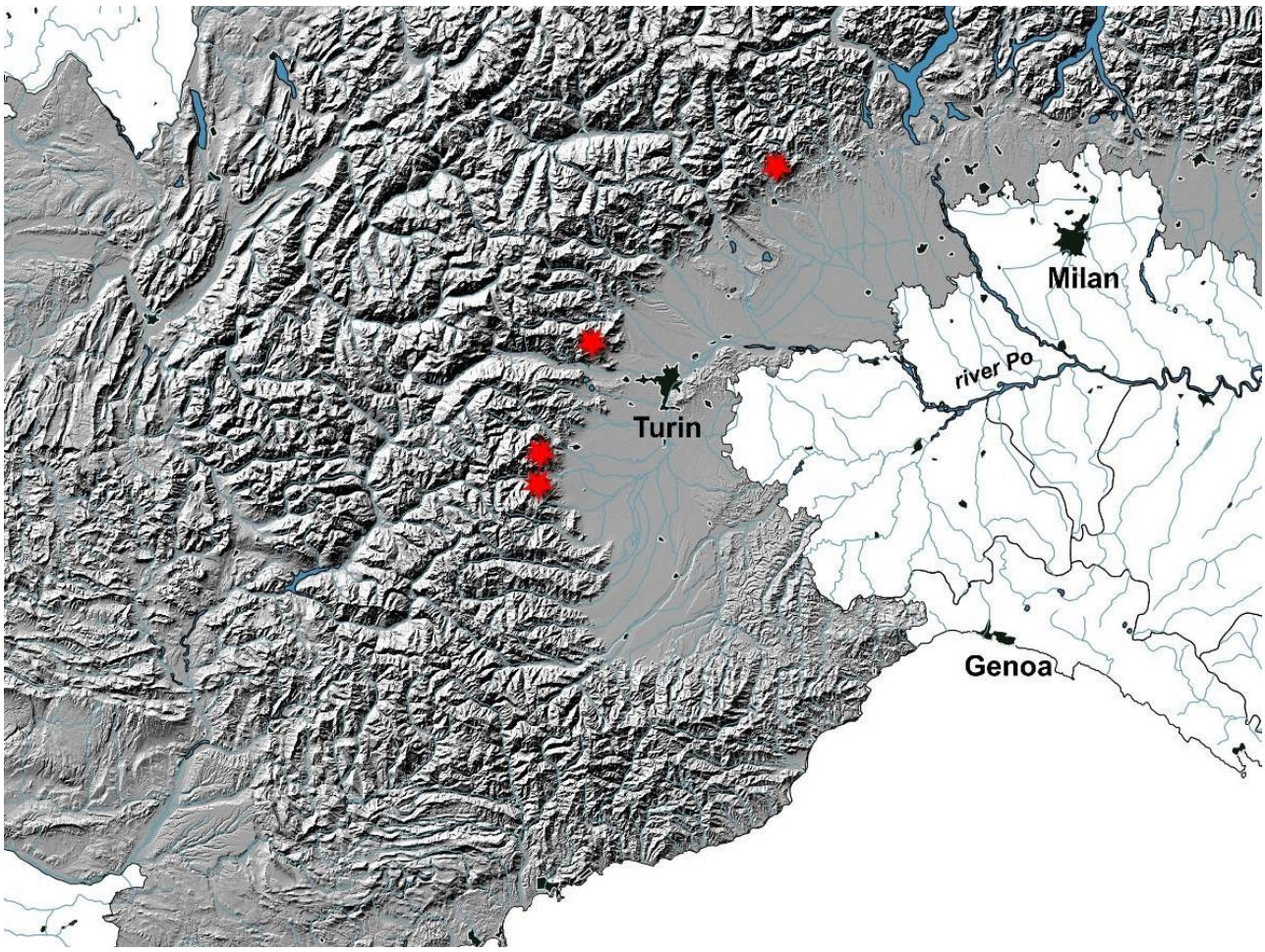

Figure S1. Map of the localities (red stars) where the *T. pyrenaeus* were collected. The area shaded in grey denotes the South-west Alpine area. Top to bottom: Bocchetto Sessera, Colle del Lys, Colle della Vaccera, and Rorà.

DNA extraction

Individuals were first rehydrated in 1 mL of a solution composed of 10 mM Tris-HCl pH 8.5, 30 mM NaCl, and 5 mM EDTA for 1 h. One leg from each individual was then used for DNA extraction. Tissues were digested overnight with proteinase K at a concentration of 10 mg/mL. Proteins were removed by adding 400 µL of 4 M sodium perchlorate followed by chloroform extraction. Total DNA

was precipitated from the upper phase by isopropylic alcohol precipitation and resuspended in TE buffer.

A fragment of mt-*CoxI* was amplified with primers based on (Folmer *et al.*, 1994), and modified as in (Astrin & Stuben, 2008): fw: LCO1490-JJ, CHACWAAYCATAAAGATATYGG; rev: HCO2198-JJ, AWACTTCVGGRTGVCCAAARAATCA.

In addition, sequences corresponding to the internal transcribed spacer 2 (ITS2) of the ribosomal DNA were amplified using primers based on (Vahtera & Muona, 2006): fw: 5'-GGGTCGATGAAGAACGCAGC-3'; rev: 5'-ATATGCTTAAATTCAGCGGG-3'.

DNA was amplificated as follows: 15 min of initial denaturation (95°C) followed by 10 cycles of 30 seconds (s) at 94°C, 45 s at 60°C down to 50°C (the annealing temperature was lowered by 1°C in each cycle), 2 min at 72°C followed by 30 cycles of 30 s at 94°C, 45 s at 50°C, 2 min at 72°C and a final extension cycle of 15 min at 72°C. All PCRs were performed in a total volume of 30 µL with HotStarTaq Master Mix (Qiagen). Both purification and sequencing were performed by an external service (Genechron, Roma). Both strands were sequenced.

Sequences were manually edited for sequential errors using MegaX software (Kumar *et al.*, 2018), then aligned using ClustalW algorithm (Thompson *et al.*, 1994). Species identity was determined based on trimmed equal length sequences by comparing each sequence against those in GenBank using the basic local alignment search tool (BLAST) function. A similarity cut-off of ≥ 97% was used for species-level identification for sequences submitted to the GenBank database.

Barcode gap distance

The record GBCL36286-19 (*COI* sequence of *T. pyrenaeus* from Spain) was obtained from BOLD Systems v4 (Ratnasingham & Hebert, 2007) and included in the FASTA matrix to evaluate the barcode gap distance among the individuals using the software Automatic Barcode Gap Discovery

(ABDG, (Puillandre *et al.*, 2012)). The default settings for Jukes-Cantor distance were applied, but the value of *X* (i.e., the proxy for the minimum gap width) was set at 1.5, 1.0 and 0.5.

The ABDG analysis did not partition the dataset with $X = 1.5$ and $X = 1.0$, thus a further run was made setting $X = 0.5$, after which two groups of sequences were obtained (group1 included the specimen from Spain only, and group2 included all Piedmont individuals), with the barcode gap distance = 0.039, and the prior maximal distance (*P*) ranging from 1.00e-3 to 7.74e-3. Each of these partitions gave the same resulting tree (Fig. S2).

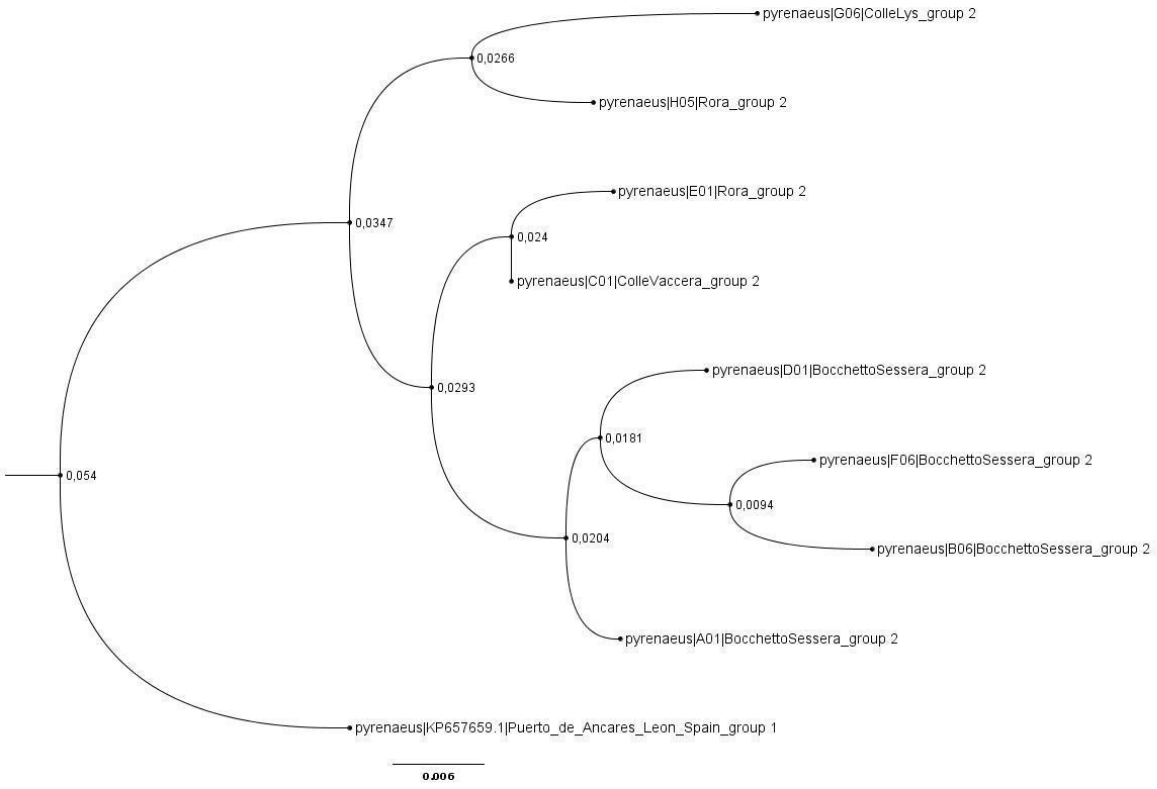

Figure S2. Resulting tree from the barcode gap analysis.

Barcode gap analysis produced a single group including the whole dataset, with $P = 1.29e\text{-}2$. The mean distance between the *T. pyrenaeus* from Spain and the individuals collected from Bocchetto Sessera in the province of Biella (BI) was 0.064; the mean distance between individuals collected from the localities in the province of Torino (TO) was 0.054; and the mean distance for all the Piedmontese specimens combined vs the Spanish one was 0.059. The mean distance between the

individuals from Bocchetto Sessera was 0.020, whereas it was 0.026 for those collected from the three localities in the province of Torino, see the distance matrix (Table S1).

pyrenaeus | Spain	0.000								
pyrenaeus | B06 | Bocchetto Sessera	0.073	0.000							
pyrenaeus | F06 | Bocchetto Sessera	0.068	0.015	0.000						
pyrenaeus | D01 | Bocchetto Sessera	0.059	0.026	0.021	0.000					
pyrenaeus | A01 | Bocchetto Sessera	0.058	0.027	0.021	0.012	0.000				
pyrenaeus | C01 | Colle Vaccera	0.049	0.029	0.028	0.023	0.015	0.000			
pyrenaeus | E01 | Rora	0.053	0.046	0.043	0.032	0.027	0.005	0.000		
pyrenaeus | H05 | Rora	0.060	0.049	0.044	0.037	0.034	0.026	0.029	0.000	
pyrenaeus | G06 | Colle Lys	0.054	0.060	0.060	0.055	0.055	0.031	0.041	0.027	0.000

Table S1. Distance matrix for the barcode gap analysis

References

Astrin, J.J. & Stüben, P.E. (2008) Phylogeny in cryptic weevils: molecules, morphology and new genera of western Palaearctic Cryptorhynchinae (Coleoptera: Curculionidae). *Invertebrate Systematics*, **22**, 503–522.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, **35**, 1547–1549.

Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. (2012) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, **21**, 1864–1877.

Ratnasingham, S. & Hebert, P.D.N. (2007) bold: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes*, **7**, 355–364.

Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.

Vahtera, V. & Muona, J. (2006) The molecular phylogeny of the Miarus campanulae (Coleoptera: Curculionidae) species group inferred from CO1 and ITS2 sequences. *Cladistics*, **22**, 222–229.