Is Acerentomon italicum Nosek, 1969 (Protura: Acerentomidae) a species complex?

Authors: Galli, Loris, Zinni, Matteo, Shrubovych, Julia, and Colasanto, Elisa

Source: Revue suisse de Zoologie, 128(1) : 121-133

Published By: Muséum d'histoire naturelle, Genève

URL: https://doi.org/10.35929/RSZ.0040

Revue suisse de Zoologie (March 2021) 128(1): 121-133 ISSN 0035-418

Is *Acerentomon italicum* Nosek, 1969 (Protura: Acerentomidae) a species complex?

Loris Galli[1*], Matteo Zinni[1], Julia Shrubovych[2,3,4] & Elisa Colasanto[1]

[1] *Dipartimento di Scienze della Terra, dell'Ambiente e della Vita – Università di Genova, Corso Europa 26, I-16132 Genova, Italy*
[2] *Institute of Systematics and Evolution of Animals, Polish Academy of Sciences, Sławkowska 17, Pl 31-016 Krakow, Poland*
[3] *State Museum of Natural History, Ukrainian Academy of Sciences, Teatral'na 18, UA 79008 Lviv, Ukraine*
[4] *Institute of Soil Biology, Biology Centre, Czech Academy of Sciences, Na Sádkách 7, 370 05 České Budějovice, Czech Republic*
* *Corresponding author: loris.galli@unige.it*

Abstract: *Acerentomon italicum* Nosek, 1969 is the most abundant Protura species in Italy. Two groups of populations characterized by the different position of a pore on tergite VII were analyzed. Specimens from Austria and from the Italian provinces of Brescia and Bolzano (Lombardy and Alto Adige, respectively) possess a postero-sub-medial pore (*psm*); in specimens from Trento (Trentino), Veneto, Friuli-Venezia Giulia and Liguria the pore is in an antero-sub-lateral (*asl*) position. An integrative approach was applied to assess other characters distinguishing these populations. The lengths of body, head, foretarsus, claw of anterior leg, foretarsal sensilla *a* and *b* were measured; in addition some ratios were calculated: TR (length of foretarsus/length of its claw), BS (distance of sensillum *t1* from base of foretarsus/distance of *t1* from foretarsus tip), length of *a*/foretarsus length, length of *b*/foretarsus length and length of *a*/length of *b*. A geometric morphometric analysis was performed on the position of setae on tergite VII, selecting as landmarks the bases of setae *A1*, *A2*, *A3*, *P1*, *P1a*, *P2*, *P2a* and *P3*. Nucleotide sequences COI-5P and 28S-D2-D3 of both groups were compared. Significant differences in the foretarsus and anterior claw length, in the arrangement of setae on tergite VII and in the barcode sequences COI-5P between the "*psm* and *asl* populations" are shown to exist. Based on the results obtained, the two groups are different, but there is not enough support to differentiate them at the species level.

Keywords: DNA barcoding - geometric morphometrics - genetic distance - integrative systematics - morphometry - porotaxy.

INTRODUCTION

Acerentomon italicum Nosek, 1969 is the most abundant Protura species in Italy (Galli *et al.*, 2016) and it is widespread in Switzerland, Austria, Corsica and Slovenia. A recent redescription of the species (Galli *et al.*, 2016) highlighted the difference between two groups of populations with regards to the porotaxy of tergite VII. Porotaxy was used as a diagnostic character to distinguish Protura species for about thirty years (e.g. Szeptycki, 1991), and it was already pointed out that the position of a single pore can differ between two closely related species. A good example is the difference in porotaxy between *Gracilentulus gracilis* Berlese, 1908b and *G. europaeus* Szeptycki, 1993 highlighted by Szeptycki (1993). Throughout most of its range (including the type area) specimens of *A. italicum* possess a single antero-sub-lateral (*asl*) pore close to the base of seta *A3*. Specimens from Switzerland, Austria, Lombardy, Friuli-Venezia Giulia and Slovenia, however, have a postero-sub-medial (*psm*) pore close to the base of seta *P1* instead. The latter is the typical position of such a pore in the Acerentomidae (Shrubovych, 2014). The remaining porotaxy and chaetotaxy of these two groups is largely identical, with a certain level of variability that characterizes all species. Moreover, DNA barcoding showed a remarkable distance between a group of three Ligurian populations (*asl*) and one from Austria (*psm*) (Galli *et al.*, 2016). These results prompted us to look for further characteristics that can distinguish the two "groups of populations" (from here on indicated as *asl* and *psm* populations). In this paper the results of an integrative study, based on genetic data, specific morphometric analysis and through a geometric morphometric analysis,

Manuscript accepted 12.01.2021
DOI: 10.35929/RSZ.0040

are shown. Geometric morphometrics, in particular, is a groundbreaking approach in Protura that deserves more attention in the future.

MATERIAL & METHODS

Morphometry

Abbreviations of depositories: MHNG = Natural History Museum of Geneva, Switzerland; UniGe = Genoa University collection, Italy; ISEAPol = Institute of Systematics and Evolution of Animals, Polish Academy of Sciences, Poland.

Material examined: MHNG; 2 ♀; Austria, Gaisberg. – MHNG; 1 ♂, 1 ♀; Austria, St Pankraz. – MHNG; 7 ♂, 7 ♀; Austria, Plabutsh. – ISEAPol; 1 ♀; Austria, Leopoldsberg, Vienna. – UniGe; 5 ♂, 13 ♀; Italy, Lombardy, Brescia Province, Polaveno. – MHNG; 1 ♂; Italy, Trentino-Alto Adige, Bolzano Province, Brunico. – MHNG; 2 ♀; Italy, Trentino-Alto Adige, Trento Province, Levico Terme. – MHNG; 5 ♂, 4 ♀; Italy, Veneto, Belluno Province, Ponte nelle Alpi. – MHNG; 3 ♂, 2 ♀; Italy, Friuli-Venezia Giulia, Udine Province, Paluzza. – MHNG; 1 ♂; Friuli-Venezia Giulia, Udine Province, surroundings of Tolmezzo. – UniGe; 1 ♂, 5 ♀; Italy, Liguria, Savona Province, Bergeggi. – UniGe; 5 ♂, 5 ♀; Italy, Liguria, Genoa Province, Genoa Quinto.

For a map of the localities listed above, see Fig. 1.

Measurements: The presence of a pore in the *asl* or *psm* position was detected (Fig. 2). When possible, the following measurements (and ratios of two measurements) of each specimen were taken: body, head, foretarsus and claw of anterior leg lengths (in the following simply referred to as tarsus and claw, as commonly used in the literature on Protura); TR (length of tarsus/length of claw) and BS (distance of sensillum *t1* from base of tarsus/distance of *t1* from tarsus tip) ratios (see Nosek, 1973); foretarsal sensilla *a* and *b* lengths; ratio of length of *a*/tarsus length, ratio of length of *b*/tarsus length and ratio of length of *a*/length of *b*.

Fig. 1. Geographical distribution of examined populations of *Acerentomon italicum* (grey = with *psm* pore; black = with *asl* pore).

Measurements were taken with the aid of an interference contrast microscope (Leica DM LB2), a Leica DFC 295 camera and a Leica Application Suite Vers. 3.8. Data were archived separately based on pore position (*asl* or *psm* populations) and sex of specimens.

Statistical analyses: For each character measured outliers were removed before the application of the following analyses. On the measurements taken One-way ANOVA and Mann-Whitney pairwise test (adopting a Bonferroni correction of p-values) were performed to see differences between sexes within the same "population" and between values recorded in *psm* and *asl* specimens within each sex. For statistically significant variables a linear correlation with the geographic coordinates of sampling localities was performed to verify any latitudinal or longitudinal trend. Analyses were made using the software "PAST" (Paleontological Statistics) version 4.02 (Hammer *et al.*, 2001).

Geometric morphometrics

Landmarks placement: The Leica system described above was also used to take photos of tergite VII of 53 specimens belonging to both populations (*asl*: 15 males and 18 females - *psm*: 9 males and 11 females). Specimens were selected in which on one side of tergite VII the landmarks mentioned below were coplanar, resulting in all being in focus at the same time. Based on these photos the coordinates of defined sets of landmarks were acquired for each specimen. The sets of coordinates (called landmark configurations) contain the information related to shape and size, orientation and location of setae on tergite VII. In particular, eight type I bidimensional landmarks (Bookstein, 1997) were defined: the bases of setae *A1*, *A2*, *A3*, *P1*, *P1a*, *P2*, *P2a* and *P3* (Fig. 2). The TPS software was used for image storing and landmarks placing (TPSutil, version 1.74, see Rohlf, 2017a and TpsDig, version 2.30, see Rohlf, 2017b).

Further analyses were performed using the "geomorph" package (Adams *et al.*, 2020), written in the "R" scientific computing language (R Core Team, 2020). Landmarks previously digitized with TPS Dig were imported into the R environment using the "readland.tps()" function.

Generalized procrustes analysis: The first step of any geometric morphometric analysis is to perform a Procrustes superimposition of the raw coordinate data (for a general overview of geometric morphometrics see Adams *et al.*, 2013). A Generalized Procrustes Analysis (GPA, see Rohlf & Slice, 1990) translates all specimens to the origin, scales them to unit-centroid size, and optimally rotates them (using a least-squares criterion) until the coordinates of corresponding landmarks align as closely as possible. The resulting aligned Procrustes coordinates represent the shape of each specimen, and are found in a curved space related to Kendall's shape space (Kendall, 1984). To perform this analysis the "gpagen()" function was used. GPA allows to extract the shape information from all landmark configurations

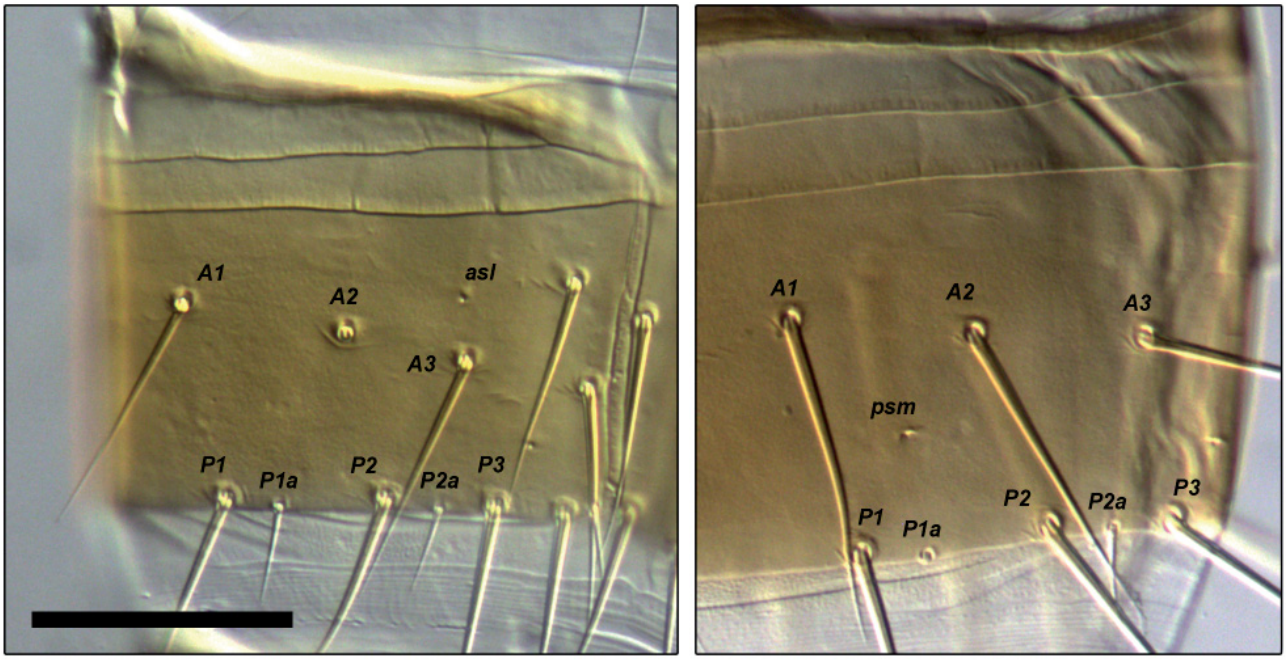

Fig. 2. *Acerentomon italicum*, tergite VII: landmarks used for geometric morphometric analysis of *asl* (left) and *psm* (right) populations. Scale bar = 50 µm.

by removing the parameters of non-shape variation (i.e. differences in size, location and orientation). Information carried by location is removed by shifting the landmarks configurations until they share the same origin in a common coordinate system (Theska *et al.*, 2020) where each landmark is represented by shape variables called "Procrustes coordinates". A data-cleaning process using the "plotOutliers()" function led to remove from the resulting analyses four specimens that proved to be outliers: one male and two females belonging to the *psm* group, and one female belonging to the *asl* group. Aligned coordinates of the remaining 49 specimens were represented using the "plotAllSpecimens()" function that creates a plot of the landmark coordinates for all specimens. This is useful for examining patterns of shape variation after GPA. Selecting "mean=TRUE", the mean shape was calculated and added to the plot (Fig. 3). Together with procrustened coordinates, the consensus and the centroid size were retrieved. Centroid size (CS) is the most common and explicit measure of size in geometric morphometrics. It is computed as the square root of the sum of the squared distances of all landmarks from their centroid (Rohlf & Slice, 1990; Goodall, 1991; Dryden & Mardia, 1998).

Principal Component Analysis: The Principal Component Analysis (PCA) is probably the most common method used to visualize the general patterns of morphological variation in multidimensional data (in this case the covariance matrix of the procrustened coordinates of landmarks). By definition, the first principal component (PC1) retains the maximum possible variance (Jolliffe, 2002; Abdi & Williams, 2010). The cumulative variance explained by each PC was examined and the relative screeplot was used to select the number of PCs for further analyses.

The Wilcoxon-Mann-Whitney test was performed to highlight differences between groups. Extreme shape variation along the PCs was explored using the "plotRefToTarget()" function. To assess the reliability of the PCA results, a cross validated (using leave-one-out jacknife procedure) linear discriminant analysis (LDA) (MASS package, Ripley *et al*., 2013) was run on extracted PCs. Discrimination was performed separately between populations and sexes.

Thus scores were considered as the projection of the centered data in the linear space defined by the eigenvectors (Jolliffe, 2002). Therefore they were analyzed to understand shape variations between groups into a dimensionally reduced space. Nevertheless the loadings were used to assess the weight of each variable in the description of the shape variation in every PC using the "fviz_contrib()" function from the "factoextra" package (Kassambara & Mundt, 2017).

Covariation method: To infer differences in shape according to factors or continuous variables, the "procD.lm()" function was used. It performs Procrustes ANOVA with permutation procedures to verify statistical hypotheses describing patterns of shape variation and covariation for a set of Procrustes-aligned coordinates. The mean shape variation among factors (populations and sexes) was tested by means of a Goodall's F test (Goodall, 1991). A second model (multivariate regression) was

used to describe shape variation according to size. Since size in geometric morphometrics is represented by centroid size, a simple multivariate regression, where shape represents the response variable and the centroid size represents the independent variable, was fitted to the data. Lastly a multivariate analysis of variance (single factor MANCOVA) was used to test changes in shape between groups, while accounting for shape varying with size.

Genetic distances

Nucleotide sequences COI-5P and 28s-D2-D3 available on the Barcode of Life Data systems (BOLD - https://www.boldsystems.org/) for *Acerentomon* species belonging to *doderoi*-group (sensu Nosek, 1973) were analyzed: *A. italicum*, *A. christiani* Shrubovych & Resh in Shrubovych *et al.*, 2016, *A. maius* Berlese, 1908b (Tables 1-2). Alignment of sequences was performed using "MUSCLE", as implemented in Mega v. 7.0.20 (Kumar *et al.*, 2016). The genetic distances based on Kimura 2-parameter (Kimura, 1980) were calculated within groups and between groups using the appropriate menu on Mega v. 7.0.20. In order to evaluate their taxonomic significance, the specimens belonging to the *A. italicum psm* population were considered as part of a separate Molecular Operational Taxonomic Unit (MOTU).

RESULTS

Morphometry

Values of the variables measured in *asl* and *psm*

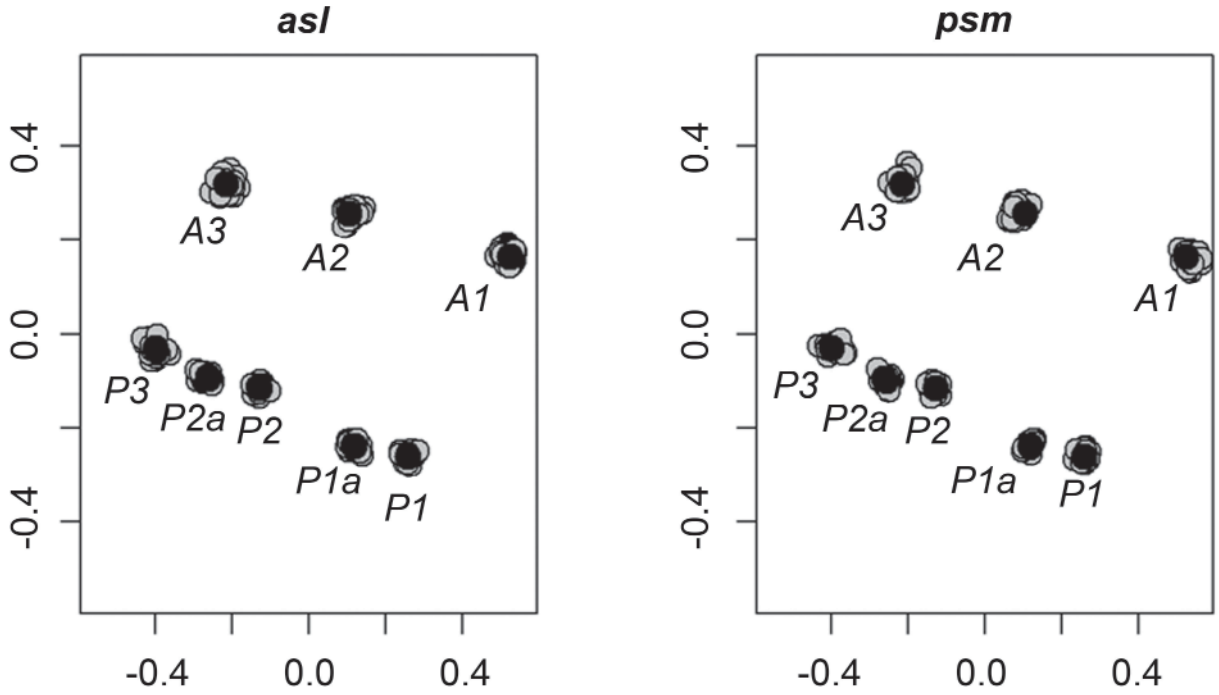

Fig. 3. Procrustes coordinates: landmarks configuration for the examined populations (*asl*, left, *psm* right). Black dots represent the mean shape (consensus).

Table 1. Accession numbers at Barcode of Life Data systems (BOLD) for specimens belonging to the *Acerentomon doderoi*-group whose nucleotide sequences COI-5P were used in this work. *Acerentomon italicum* specimens without notes belong to Italian populations with an *asl* pore.

BOLD Nr|Species|Sequence analysed	NOTES
PROAT001-12|Acerentomon_maius|COI-5P	
PROAT005-12|Acerentomon_maius|COI-5P	
PROAT006-12|Acerentomon_maius|COI-5P	
PROAT015-12|Acerentomon_maius|COI-5P	
PROAT016-12|Acerentomon_maius|COI-5P	
PROAT017-12|Acerentomon_maius|COI-5P	
PROAT018-12|Acerentomon_maius|COI-5P	
PROAT019-12|Acerentomon_maius|COI-5P	

Table 1 (continued)

BOLD Nr|Species|Sequence analysed	NOTES
PROAT020-12|Acerentomon_maius|COI-5P	
PROAT024-12|Acerentomon_maius|COI-5P	
PROAT025-12|Acerentomon_sp._gr_doderoi|COI-5P	*Acerentomon christiani*, see Shrubovych *et al.*, 2016
PROAT026-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT027-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT029-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT030-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT031-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT032-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT033-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT034-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT035-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT036-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT037-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT038-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT039-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT040-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT041-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT042-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT054-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT055-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT063-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT064-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT065-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT066-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT069-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT071-12|Acerentomon_sp._gr_doderoi|COI-5P	
PROAT052-12|Acerentomon_italicum|COI-5P	From Leopoldsberg (AUT), with *psm* pore
PROIT001-15|Acerentomon_italicum|COI-5P	
PROIT002-15|Acerentomon_italicum|COI-5P	
PROIT003-15|Acerentomon_italicum|COI-5P	
PROIT004-15|Acerentomon_italicum|COI-5P	
PROIT005-15|Acerentomon_italicum|COI-5P	
PROIT006-15|Acerentomon_italicum|COI-5P	
PROIT007-15|Acerentomon_italicum|COI-5P	
PROIT008-15|Acerentomon_italicum|COI-5P	
PROIT009-15|Acerentomon_italicum|COI-5P	
PROIT010-15|Acerentomon_italicum|COI-5P	
PROIT011-15|Acerentomon_italicum|COI-5P	
PROIT012-15|Acerentomon_italicum|COI-5P	
PROIT013-15|Acerentomon_italicum|COI-5P	
PROIT014-15|Acerentomon_italicum|COI-5P	
PROIT015-15|Acerentomon_italicum|COI-5P	
PROIT016-15|Acerentomon_italicum|COI-5P	
PROIT017-15|Acerentomon_italicum|COI-5P	
PROIT018-15|Acerentomon_italicum|COI-5P	
PROIT019-15|Acerentomon_italicum|COI-5P	
PROIT020-15|Acerentomon_italicum|COI-5P	
PROIT021-15|Acerentomon_italicum|COI-5P	

Table 2. Accession numbers at Barcode of Life Data systems (BOLD) for specimens belonging to the *Acerentomon doderoi*-group whose nucleotide sequences 28S-D2-D3 were used in this work. *Acerentomon italicum* specimens without notes belong to Italian populations with an *asl* pore.

BOLD Nr|Species|Sequence analysed	**NOTES**
PROAT001-12|Acerentomon_maius|28S-D2-D3	
PROAT006-12|Acerentomon_maius|28S-D2-D3	
PROAT020-12|Acerentomon_maius|28S-D2-D3	
PROAT005-12|Acerentomon_maius|28S-D2-D3	
PROAT015-12|Acerentomon_maius|28S-D2-D3	
PROAT016-12|Acerentomon_maius|28S-D2-D3	
PROAT018-12|Acerentomon_maius|28S-D2-D3	
PROAT024-12|Acerentomon_maius|28S-D2-D3	
PROAT096-13|Acerentomon_maius|28S-D2-D3	
PROAT017-12|Acerentomon_maius|28S-D2-D3	
PROAT019-12|Acerentomon_maius|28S-D2-D3	
PROAT027-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	*Acerentomon christiani*, see Shrubovych *et al*., 2016
PROAT030-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT033-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT037-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT040-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT042-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT063-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT031-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT032-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT034-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT035-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT069-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT071-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT100-13|Acerentomon_sp._gr._doderoi|28S-D2-D3	
PROAT026-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT036-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT041-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT054-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT055-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT064-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT065-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT066-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT038-12|Acerentomon_sp._gr_doderoi|28S-D2-D3	
PROAT092-13|Acerentomon_sp.|28S-D2-D3	*A. italicum* Austria, with *psm* pore
PROAT091-13|Acerentomon_cf._italicum|28S-D2-D3	*A. italicum* Austria, with *psm* pore
PROAT052-12|Acerentomon_italicum|28S-D2-D3	*A. italicum* Austria, with *psm* pore
PROIT003-15|Acerentomon_italicum|28S-D2-D3	
PROIT007-15|Acerentomon_italicum|28S-D2-D3	
PROIT009-15|Acerentomon_italicum|28S-D2-D3	
PROIT011-15|Acerentomon_italicum|28S-D2-D3	
PROIT014-15|Acerentomon_italicum|28S-D2-D3	
PROIT016-15|Acerentomon_italicum|28S-D2-D3	
PROIT020-15|Acerentomon_italicum|28S-D2-D3	
PROIT004-15|Acerentomon_italicum|28S-D2-D3	
PROIT006-15|Acerentomon_italicum|28S-D2-D3	

Table 2 (continued)

BOLD Nr|Species|Sequence analysed	NOTES
PROIT010-15|Acerentomon_italicum|28S-D2-D3	
PROIT015-15|Acerentomon_italicum|28S-D2-D3	
PROIT017-15|Acerentomon_italicum|28S-D2-D3	
PROIT019-15|Acerentomon_italicum|28S-D2-D3	
PROIT001-15|Acerentomon_italicum|28S-D2-D3	
PROIT002-15|Acerentomon_italicum|28S-D2-D3	
PROIT008-15|Acerentomon_italicum|28S-D2-D3	
PROIT012-15|Acerentomon_italicum|28S-D2-D3	
PROIT013-15|Acerentomon_italicum|28S-D2-D3	
PROIT018-15|Acerentomon_italicum|28S-D2-D3	
PROIT021-15|Acerentomon_italicum|28S-D2-D3	

Table 3. *Acerentomon italicum*: morphometry of *asl* and *psm* populations. Mean values ± standard deviations; lengths in µm.

	psm		***asl***	
	Males (n = 14)	**Females (n = 24)**	**Males (n = 15)**	**Females (n = 18)**
Body length	1600 ± 142	1696 ± 190	1533 ± 91	1679 ± 76
Head length	214 ± 15	224 ± 11	203 ± 12	208 ± 15
Tarsus length	114 ± 7	118 ± 4	104 ± 4	109 ± 5
Claw length	41 ± 1	44 ± 3	40 ± 2	40 ± 2
TR	2.53 ± 0.13	2.61 ± 0.14	2.48 ± 0.14	2.58 ± 0.11
BS	0.57 ± 0.04	0.56 ± 0.06	0.54 ± 0.04	0.54 ± 0.03
Sensillum *a* length	30 ± 2	31 ± 2	29 ± 2	30 ± 2
Sensillum *b* length	19 ± 2	20 ± 2	18 ± 1	19 ± 2
Sensillum *a*/tarsus	0.25 ± 0.01	0.25 ± 0.02	0.27 ± 0.02	0.27 ± 0.02
Sensillum *b*/tarsus	0.16 ± 0.01	0.16 ± 0.02	0.17 ± 0.01	0.16 ± 0.01
Sensillum *a*/Sensillum *b*	1.49 ± 0.15	1.52 ± 0.16	1.53 ± 0.11	1.58 ± 0.15

populations are shown in Table 3. The One-way ANOVA and Mann-Whitney pairwise test (adopting a Bonferroni correction of p-values) showed only two variables with statistically significant differences between *psm* and *asl* populations: tarsus length ($F = 23.22$; $p << 0.001$) and claw length ($F = 12.18$; $p = 0.0025$). Both characters have significantly larger values in *psm* than *asl* populations (Table 4, Fig. 4). For none of the characters it was possible to outline a latitudinal or longitudinal trend. Larger sizes of females than males within the same group were not statistically significant.

Table 4. Mann-Whitney pairwise test (adopting Bonferroni correction of p-values). Output for tarsus length and claw length.

Tarsus length				
	***psm* M**	***psm* F**	***asl* M**	***asl* F**
***psm* M**		1	< 0.001	0.044
***psm* F**	1		<< 0.001	<< 0.001
***asl* M**	< 0.001	<< 0.001		0.1884
***asl* F**	0.044	<< 0.001	0.1884	
Claw length				
	***psm* M**	***psm* F**	***asl* M**	***asl* F**
***psm* M**		0.85	0.017	0.005
***psm* F**	0.85		0.004	< 0.001
***asl* M**	0.017	0.004		1
***asl* F**	0.005	< 0.001	1	

Geometric morphometrics

The first four components of PCA accounted for more than 73% of the total shape variation (Table 5). Also the screeplot inspection (Fig. 5) confirmed the same number of principal components as a reliable measure. The scoreplot based on the first two PCs shows that the two populations were quite different in shape (Fig. 6).

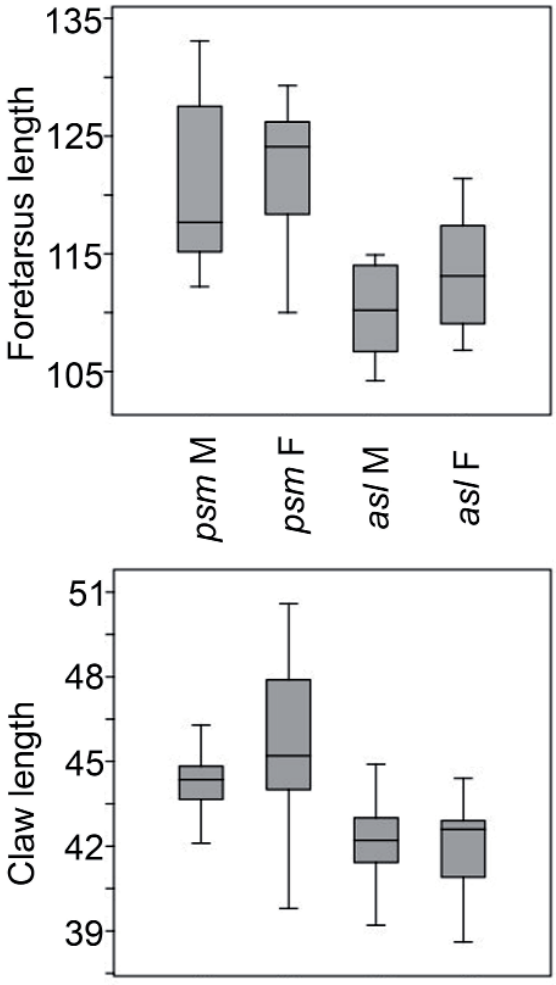

Fig. 4. Boxplot of tarsus and claw length measurements in both sexes of *psm* and *asl* populations of *Acerentomon italicum*.

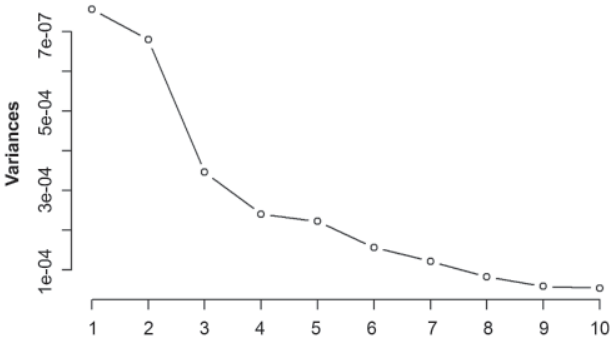

Fig. 5. Screeplot of PCA run on procrustened coordinates.

Table 5. Results of principal components analysis (PCA). For each PC the amount of explained variance is reported together with the cumulative proportions (both in %).

	Variance	Cumulative
PC01	27.24	27.24
PC02	24.52	51.76
PC03	12.47	64.23
PC04	8.67	72.90
PC05	8.02	80.91
PC06	5.67	86.58
PC07	4.39	90.97
PC08	2.99	93.95
PC09	2.14	96.09
PC10	1.95	98.04
PC11	1.32	99.36
PC12	0.64	100.00
PC13	0.00	100.00
PC14	0.00	100.00
PC15	0.00	100.00
PC16	0.00	100.00

Only for the first two PCs Wilcoxon-Mann-Whitney tests on differences between groups retrieved significant differences (PC1: U = 474, p = 5.83E-06; PC2: U = 360, p = 6.57E-02). When sexes were compared, no clear distinctions arose from between-group inference. Since only PC1 and PC2 accounted for shape variation, extreme shape deformations along these axes were plotted as a thin-plate spline deformation grid (Fig. 7). More than 40% of variation in PC1 was due to the horizontal coordinate of seta *A2* (*A2x*), while PC2 shows a more even situation where *P3* and *A3* accounted for most of the variation (Fig. 8).

A linear discriminant analysis performed between populations confirmed the reliability of PCA, showing that the first four PCs have an overall classification accuracy of 83.7% (41 specimens correctly classified from a total of 49, Table 6), while discrimination between males and females did not provide meaningful results (51% overall accuracy).

The Goodall test retrieved significant differences in setae arrangement only between populations (F = 7.2242, p = 1e-04 with 1 and 48 df) but not between sexes (F = 0.7429, p = 6.17E-01 with 1 and 48 df). Despite centroid size showing a strong group structure both between populations (W = 35, p-value = 2.36E-08) and sexes (W = 412, p-value = 2.32E-02), multivariate regression highlighted a barely significant amount of covariation between shape and size (F = 1.91; p = 0.05 with 1 and 48 df).

Genetic distances

Values of Kimura 2-parameter within-group and between-group genetic distances (Kimura, 1980) calculated on the basis of the nucleotide sequences COI-5P and 28s-D2-D3 available at BOLD for *Acerentomon* species belonging to the *doderoi*-group are shown in Tables 7-8. With regard to the barcode sequence COI-5P, the distance of 0.292 between the only available specimen

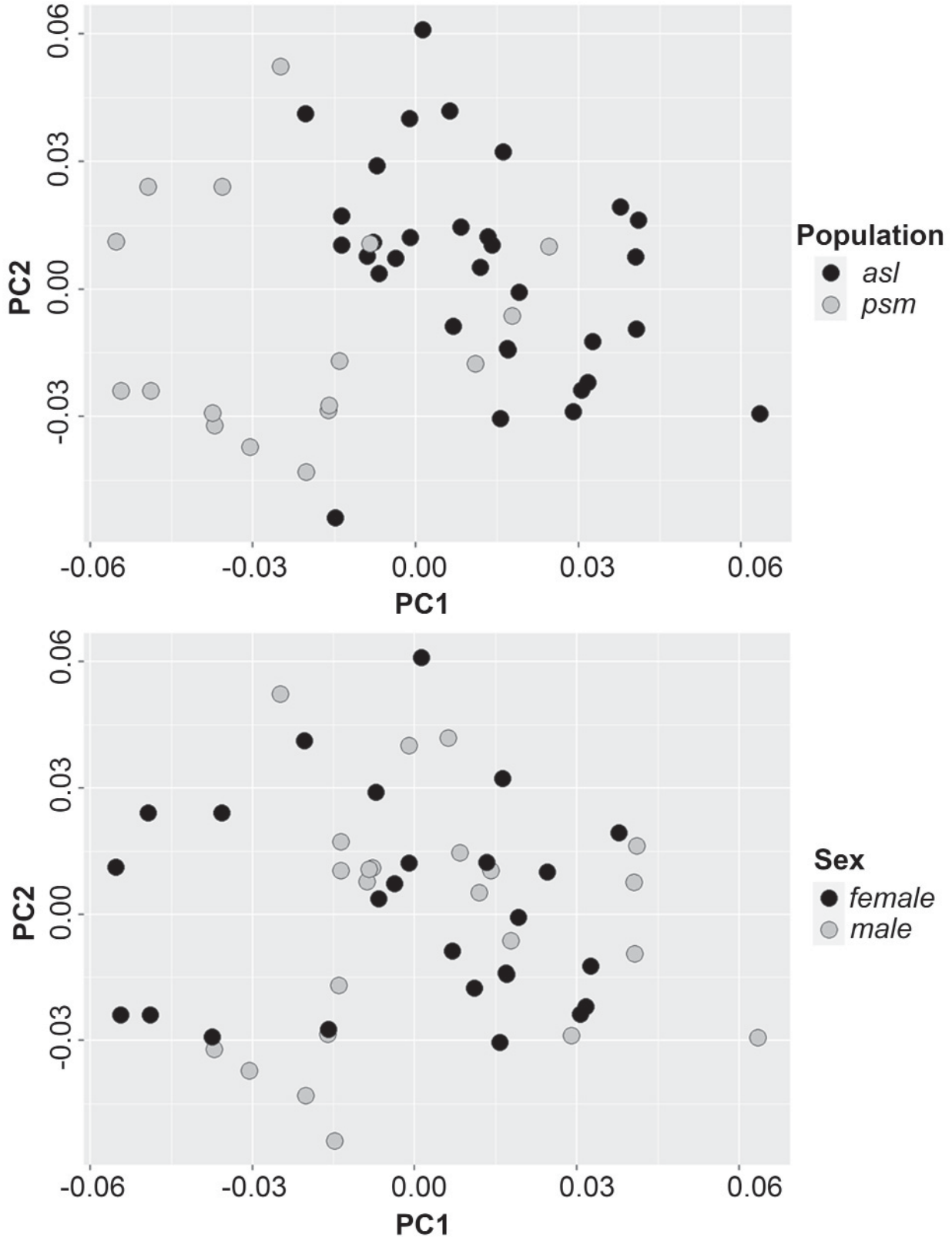

Fig. 6. Scoreplots of the principal component analysis performed on 49 *A. italicum* specimens and 8 landmarks along the first two principal axes explaining 27.24% (PC1) and 24.17% (PC2) of the shape variation.

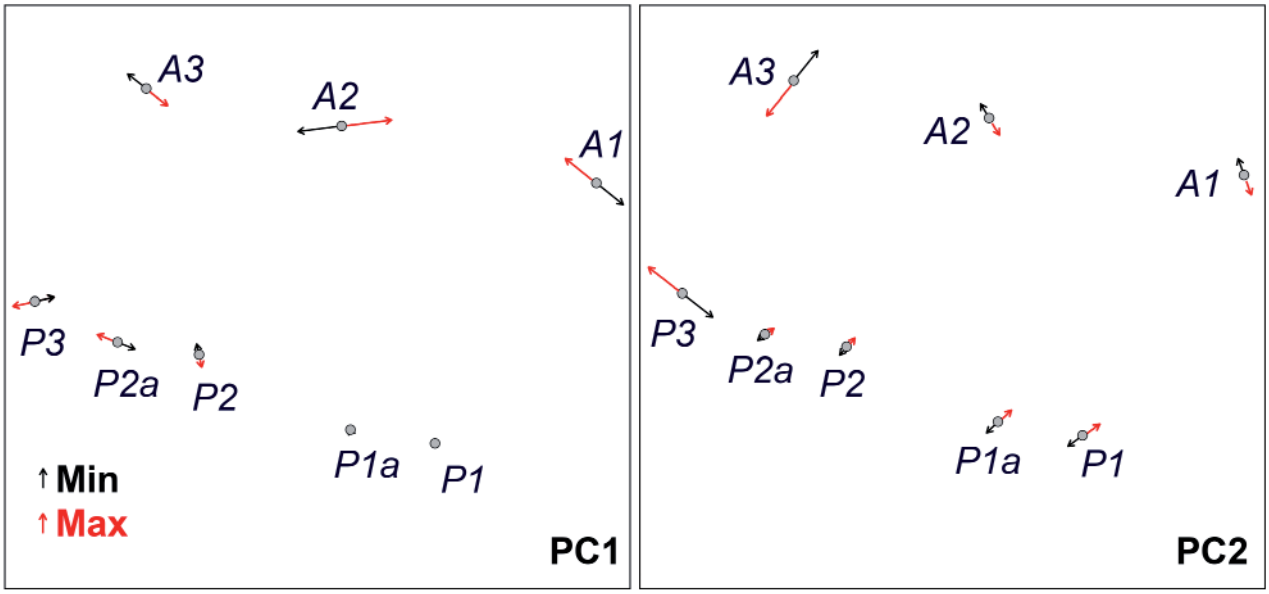

Fig. 7. Extreme shape deformation for PC1 and PC2 maximum and minimum values. Arrows show deformations from mean shape (consensus).

Table 6. Confusion matrix obtained performing LDA on first four PCs. The numbers of specimens correctly identified as belonging to *asl* (top left) and *psm* (down right) populations are given.

	asl	***psm***
asl	30	2
psm	6	11

Table 7. Genetic distances based on Kimura 2-parameter within *Acerentomon* species belonging to the *doderoi*-group (comparing COI-5P and 28S-D2-D3 sequences).

	COI-5P	28S-D2-D3
A. maius	0.002	0.001
A. christiani	0.001	0
A. italicum (psm)	-	0
A. italicum (asl)	0.014	0.075

of *A. italicum* belonging to the *psm* population and those with an *asl* pore is much higher than any within-group distance (Table 7) and it is comparable with between-group distances (Table 8). On the other hand, in the case of expansion segments 28s-D2-D3, the distance between the group of three *A. italicum* specimens with a *psm* pore (from Austria) and those with an *asl* pore is 0.077 (Table 8). This value is close to the within-group distance calculated for *asl* specimens (0.075 in Table 7) and lower than the other between-group distances (Table 8).

Table 8. Genetic distances based on Kimura 2-parameter between *Acerentomon* species belonging to the *doderoi*-group (comparing COI-5P and 28S-D2-D3 sequences).

COI-5P	*A. maius*	*A. christiani*	*A. italicum (asl)*
A. christiani	0.288		
A. italicum (asl)	0.362	0.347	
A. italicum (psm)	0.324	0.305	0.292
28S-D2-D3			
A. christiani	0.158		
A. italicum (asl)	0.168	0.218	
A. italicum (psm)	0.139	0.182	0.077

DISCUSSION

The traditional morphometric analyses for this paper were limited to few characters. We focused on those characters that, based on a recent redescription of *A. italicum* (Galli *et al.*, 2016) and on our experience, seemed to be the most promising in order to highlight differences between *psm* and *asl* populations. In particular two foretarsal sensilla (*a* and *b*) are among the most important characters for *Acerentomon* species identification (see Shrubovych *et al.*, 2016), especially in those species assumed to be close to *A. italicum* (see Galli *et al.*, 2016). Therefore the absolute and relative length of these sensilla (to the tarsus and between them) was added to the structures that are traditionally measured. However, their between-group differences are not statistically significant. Just two size variables (tarsus and claw length) differ significantly between the two groups examined. Even though we did

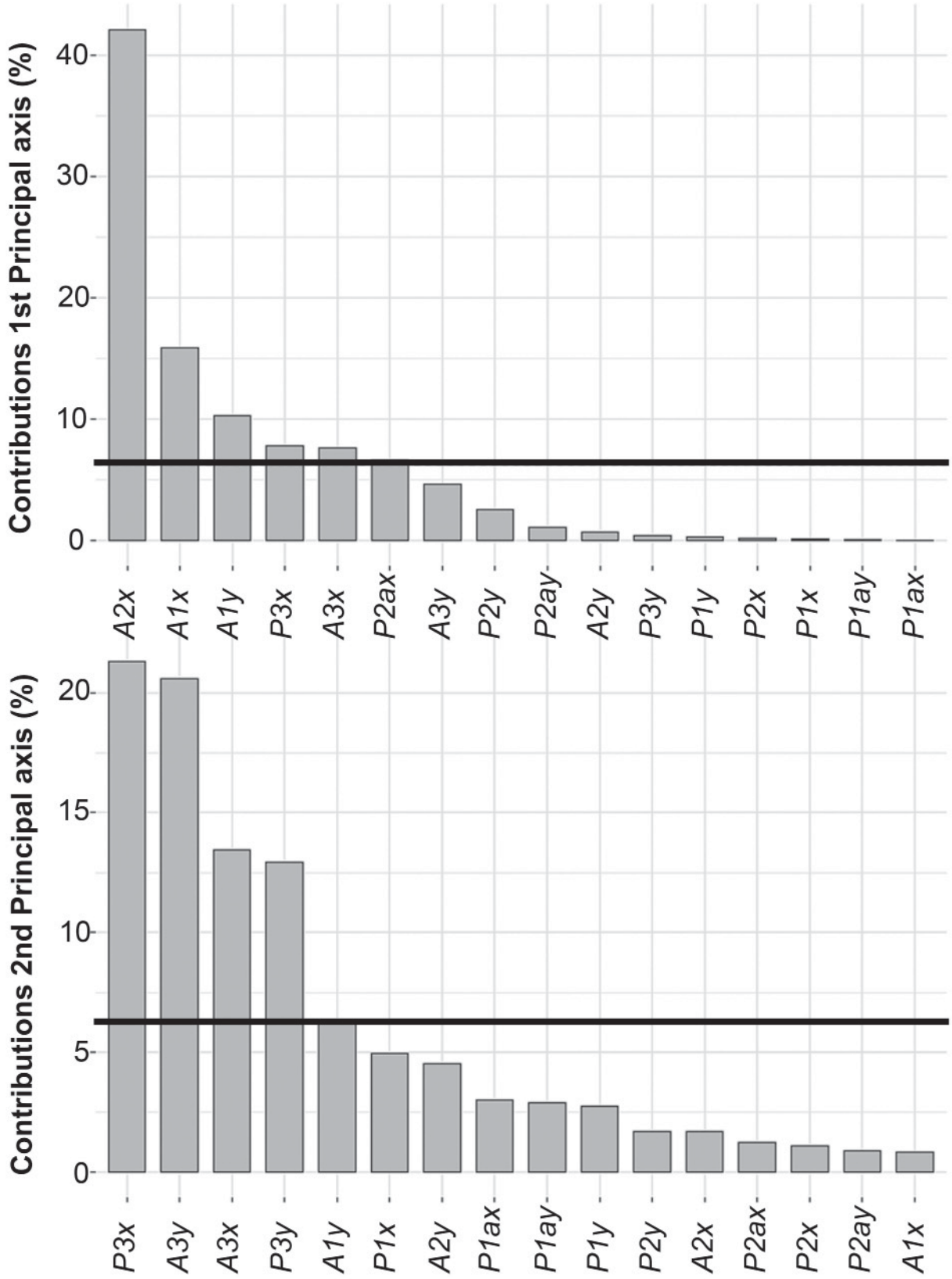

Fig. 8. Barplot showing the contribution of variables (expressed as percentage) along the first and second principal axes. The thick reference lines correspond to the expected value if the contributions were uniform. On the abscissa are the x and y coordinates of selected landmarks (setae *A1* to *A3* and *P1* to *P3*, see Fig. 2).

not outline any geographical trend for the examined variables, we cannot exclude that this may be due to the small sample examined. This could also be the reason for not having found differences in morphological characters which are of diagnostic importance in Acerentomidae (see Shrubovych, 2014).

Generally geometric morphometric analyses were applied to medium- and large-sized animals, and in hexapods they often dealt with wing venation. In the extensive literature dealing with this method, only few papers are about microscopic animals or about chaetotaxy (Mutanen *et al.*, 2009; Ramírez-Sánchez *et al.*, 2016; Theska *et al.*, 2020). This study therefore is the first one in which a geometric morphometric approach is applied to Protura and one of the first ever to deal with setae arrangement and microscopic animals. Geometric morphometrics allows to analyze shapes without having almost any information about size. This approach avoids potential bias generated by intra- and even inter-population variation that can be caused by sexual dimorphism, which is quite common among arthropods. Both explorative (PCA) and inferential (ANOVA) techniques highlight differences between the arrangement of setae on tergite VII in the two groups examined. Results from PCA showed that the largest variation is found in the arrangement of setae *A2*, *A3* and *P3*. Shape-related variation apparently does not depend on the variation of size-related elements as shown by regressions. When groups were compared, significant differences between populations were clearly highlighted by the Goodall test.

Both morphometry and geometric morphometrics do not show significant differences between sexes within each group. This confirms what was already highlighted by Galli *et al.* (2016) who examined all morphometric characters of *A. italicum* traditionally used in Protura taxonomy. However, Table 3 shows that females are on average slightly bigger than males. This size difference was evidenced also in other species of Acerentomidae like *Acerentulus shrubovychae* Galli & Capurro, 2013 and *Acerentulus tortii* Galli *et al.*, 2017.

The existence of species complexes in Protura was already suggested by Fratello & Sabatini (1989) on the basis of differences between kariotypes of different *Eosentomon transitorium* Berlese, 1908a populations. A species-level diversification can also be assumed for allopatric populations of some species that were likely separated by ancient vicariance events, as it is the case in the amphi-Atlantic *Delamarentulus tristani* (Silvestri, 1938) (see Galli & Rellini, 2020). Theoretically DNA barcoding could be an excellent tool to distinguish cryptic species, and in some cases it was used for that purpose since the early days (e.g. Herbert *et al.*, 2004). On the other hand, it was soon shown that especially for insufficiently sampled taxa (such as Protura) barcoding performs poorly when used for species identification (e.g. Meyer & Paulay, 2005). Our results must therefore be considered with extreme caution, especially in relation to the conflicting data from the analysis of COI-5P and 28S-D2-D3 sequences. For the latter some studies show that the more variable D2 region is more suitable not only for inferring relationships but also for taxonomic and diagnostic purposes at species level (e.g. Schmidt *et al.*, 2006).

CONCLUSIONS

Our integrative analyses show that *psm* and *asl* populations of *A. italicum* differ for some characters other than the porotaxy of tergite VII. In particular *psm* specimens are characterized by significantly longer foretarsi and claws. The geometric morphometric analysis of the arrangement of setae on tergite VII allows to discriminate between these populations both in descriptive and inferential terms, leaving no doubt about the shape difference between *asl* and *psm* groups. Their DNA barcode sequences COI-5P differ at an interspecific level. These data show a clear difference between the studied groups. However, it is worth pointing out that only a small sample from the available localities has been studied, with few specimens selected for each. It would be better to include more localities and more specimens. Moreover, only a single DNA sequence was available for the *psm* group. At present there is therefore not enough evidence to differentiate the two groups at the species level with reasonable certainty.

Nevertheless this study is the first one in which a geometric morphometric approach was applied to Protura, and one of the first ever to deal with the arrangement of setae and microscopic animals. Beyond the result in itself, it is important to show that this methodology can be applied to the study of evolutionary patterns and trends of chaeto- and porotaxy in Protura in addition to the traditional morphometric approach and to the genetic one. This is particularly useful when genetic data indicate significant differences between populations which supposedly belong to the same species.

ACKNOWLEDGEMENTS

We wish to thank Dr Peter Schwendinger of the Natural History Museum of Geneva for kindly providing us the specimens we analyzed for this paper. Special thanks go to Dr Yun Bu for valuable suggestions that helped greatly improving the manuscript.

REFERENCES

Abdi H., Williams L.J. 2010. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics* 2(4): 433-459.

Adams D., Rohlf F., Slice D. 2013. A field comes of age: geometric morphometrics in the 21st century. *Hystrix, the Italian Journal of Mammology* 24(1): 7-14.

Adams D., Collyer M., Kaliontzopoulou A. 2020. “Geomorph:

Software for geometric morphometric analyses. R package version 3.2.1. Available at https://cran.r-project.org/package=geomorph

Berlese A. 1908a. Nuovi Acerentomidi. *Redia* 5: 16-19, pl. I.

Berlese A. 1908b. Osservazioni intorno agli Acerentomidi. Nota preventiva. *Redia* 5: 110-122.

Bookstein F.L. 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Medical Image Analysis* 1(3): 225-243.

Dryden I.L., Mardia K.V. 1998. Statistical shape analysis. *Wiley*, *Chichester,* xvii+347 pp.

Fratello B., Sabatini M.A. 1989. Chromosome studies in Protura Eosentomoidea (pp. 167-170). *In:* Dallai R. (ed.). 3rd International Seminar on Apterygota. *University of Siena*, Siena, 489 pp.

Galli L., Capurro M. 2013. *Acerentulus shrubovychae* sp. nov. from Italy (Protura: Acerentomidae). *Zootaxa* 3609 (4): 431-436.

Galli L., Rellini I. 2020. The geographic distribution of Protura (Arthropoda: Hexapoda): a review. *Biogeographia – The Journal of Integrative Biogeography* 35: 51-69.

Galli L., Bartel D., Capurro M., Pass G., Sarà A., Shrubovych J., Szucsich N. 2016. Redescription and review of the most abundant conehead in Italy: *Acerentomon italicum* Nosek, 1969 (Protura: Acerentomidae). *Italian Journal of Zoology* 83(1): 43-58.

Galli L., Capurro M., Lionetti G., Zinni M. 2017. *Acerentulus tortii* sp. nov. from Greece (Protura: Acerentomidae). *Zootaxa* 4232(3): 437-443.

Goodall C. 1991. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society* Series B (Methodological) 53(2): 285-321.

Hammer Ø., Harper D.A.T., Ryan P.D. 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* 4: 1-9. Available at http://folk.uio.no/ohammer/past

Herbert P.D.N., Penton E.H., Burns J.M., Janzen D.H., Hallwachs W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences 101(41): 14812-14817.

Jolliffe I.T. 2002. Principal component analysis (2nd ed.). *Springer*, *New York,* 516 pp.

Kassambara A., Mundt F. 2017. Package 'factoextra'. Extract and visualize the results of multivariate data analyses, 76. Available at http://www.sthda.com/english/rpkgs/factoextra

Kendall D.G. 1984. Shape manifolds, Procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society* 16: 81-121.

Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16: 111-120.

Kumar S., Stecher G., Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33: 1870-1874.

Meyer C.P., Paulay G. 2005. DNA Barcoding: error rates based on comprehensive sampling. *PLoS Biology* 3(12): e422.

Mutanen M., Ruotsalainen H., Kaila L. 2009. Variation in larval chaetotaxy in *Orthosia gothica* (Lepidoptera: Noctuidae): effects of mother, sex and side, and implication for systematics. *Systematic Entomology* 34: 712-723.

Nosek J. 1969. Three new species of Protura from Italy. *Atti dell'Istituto Veneto di Scienze, Lettere ed Arti. Classe di Scienze matematiche e naturali* 127: 485-494.

Nosek J. 1973. European Protura. *Muséum d'Histoire naturelle, Genève,* 345 pp.

R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at https://www.R-project.org/

Ramírez-Sánchez M.M., De Luna E., Cramer C. 2016. Geometric and traditional morphometrics for the assessment of character state identity: multivariate statistical analyses of character variation in the genus *Arrenurus* (Acari, Hydrachnidia, Arrenuridae). *Zoological Journal of the Linnean Society* 177(4): 720-749.

Ripley B., Venables B., Bates D.M., Hornik K., Gebhardt A., Firth D., Ripley M.B. 2013. Package 'mass'. Cran R, 538. Available at https://cran.r-project.org/web/packages/MASS/index.html

Rohlf F.J. 2017a. tpsUtil32. New York: Department of Ecology and Evolution, State University of New York at Stony Brook. Available at http://life.bio.sunysb.edu/morph/

Rohlf F.J. 2017b. TpsDig, version 2.30. New York: Department of Ecology and Evolution, State University of New York at Stony Brook. Available at http://life.bio.sunysb.edu/morph/

Rohlf F.J., Slice D. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Biology* 39(1): 40-59.

Schmidt S., Driver F., De Barro P. 2006. The phylogenetic characteristics of three different 28S rRNA gene regions in *Encarsia* (Insecta, Hymenoptera, Aphelinidae). *Organisms, Diversity & Evolution* 6: 127-139.

Shrubovych J. 2014. Identification and character analysis of the Acerentomidae (Protura) of the northeastern Palearctic (Protura: Acerentomidae). *Zootaxa* 3755(2): 136-164.

Shrubovych J., Bartel D., Szucsich N.U., Resch M.C., Pass G. 2016. Morphological and genetic analysis of the *Acerentomon doderoi* group (Protura: Acerentomidae) with description of *A. christiani* sp. nov. *PLoS ONE* 11(4): e0148033.

Silvestri F. 1938. Primo contributo alla conoscenza dei Protura (Insecta) del Brasile e di Costa Rica (pp. 441-445). *In:* Neiva A. (ed.). Livro Jubilar do Professor Lauro Travassos. Editato para commemorar o 25º anniversario de suas actividades scientificas (1913-1938). *Instituto Oswaldo Cruz, Rio de Janeiro*, 589 pp.

Szeptycki A. 1991. Polish Protura V. Genus *Acerentulus* Berlese, 1908 (Acerentomidae). *Acta Zoologica Cracoviensia* 34(1): 1-64.

Szeptycki A. 1993. *Gracilentulus* species of "*gracilis*" group (Protura, Berberentomidae). *Acta Zoologica Cracoviensia* 35(3): 381-411.

Theska T., Sieriebriennikov B., Wighard S.S., Werner M.S., Sommer R.J. 2020. Geometric morphometrics of microscopic animals as exemplified by model nematodes. *Nature Protocols* 15(8): 2611-2644.