Turning one into five: Integrative taxonomy uncovers complex evolution of cryptic species in the harvester ant *Messor* “*structor*”

Florian M. Steiner¹,a, Sándor Csőszb, Bálint Markóc, Alexander Gamischd, Lukas Rinnhofera, Clemens Folterbauer³, Sarina Hammerle³, Christian Stauffer⁵, Wolfgang Arthofer³, Birgit C. Schlick-Steiner⁷

¹ Molecular Ecology Group, Department of Ecology, University of Innsbruck, Technikerstraße 25, 6020 Innsbruck, Austria
² Department of Biosciences, University of Salzburg, Hellbrunnerstraße 34, 5020 Salzburg, Austria
³ Hungarian Department of Biology and Ecology, Babeș-Bolyai University, Clinicilor st. 5–7, 400006 Cluj-Napoca, Romania
⁴ Department of Forest and Soil Sciences, Boku, University of Natural Resources and Life Sciences, Vienna, Peter-Jordan-Straße 82/I, 1190 Vienna, Austria
⁵ Molecular Phylogenetics and Evolution Group, Department of Ecology, University of Innsbruck, Technikerstraße 25, 6020 Innsbruck, Austria
⁶ MTA-ELTE-MTM, Ecology Research Group, Pázmány Péter sétány 1C, H-1117 Budapest, Hungary
⁷ Department of Forest and Soil Sciences, Boku, University of Natural Resources and Life Sciences, Vienna, Peter-Jordan-Straße 82/I, 1190 Vienna, Austria

A R T I C L E   I N F O

Keywords:
- Amplified fragment length polymorphism
- Ecological niche modelling
- Mitochondrial DNA
- Species delimitation
- Traditional morphometrics
- Wolbachia endosymbionts

A B S T R A C T

Seed harvesting ants are ecosystem engineers that shape vegetation, nutrient cycles, and microclimate. Progress in ecological research is, however, slowed down by poor species delimitation. For example, it has not been resolved to date, how many species the European harvester ant *Messor* “*structor*” (Latreille, 1798) represents. Since its first description, splitting into additional taxa was often proposed but not accepted later on due to inconsistent support from morphology and ecology. Here, we took an iterative integrative-taxonomy approach – comparing multiple, independent data sets of the same sample – and used traditional morphometrics, Wolbachia symbionts, mitochondrial DNA, amplified fragment length polymorphism, and ecological niche modelling. Using the complementarity of the data sets applied, we resolved multiple, strong disagreements over the number of species, ranging from four to ten, and the allocation of individuals to species. We considered most plausible a five-species hypothesis and conclude the taxonomic odyssey by redescribing *Messor* *m. structor*, *M. ibericus* Santschi, 1925, and *M. muticus* (Nylander, 1849) stat.rev., and by describing two new species, *M. ponticus* sp.n. and *M. mcarthuri* sp.n. The evolutionary explanations invoked in resolving the various data conflicts include pronounced morphological crypsis, incomplete lineage-sorting or ongoing copeciation of endosymbionts, and peripatric speciation – these ants’ significance to evolutionary biology parallels that to ecology. The successful solution of this particular problem illustrates the usefulness of the integrative approach to other systematic problems of comparable complexity and the importance of understanding evolution to drawing correct conclusions on species’ attributes, including their ecology and biogeography.

1. Introduction

Ants are major ecological players (Ward, 2006; Del Toro et al., 2012), their impact exceeding by far (Wilson and Hölldobler, 2005) the one percent of biodiversity they represent (Costello et al., 2013; Bolton, 2016). The ants of Central Europe are among those with the longest research history (Linnaeus, 1758), and they may be the best investigated among all ant faunas (Seifert, 1999). Nevertheless, Central European ants have continued experiencing taxonomic change, including very recently – some species were sunk into junior synonymy (e.g., Schlick-Steiner et al., 2005a; Csősz and Schulz, 2010; Csősz, 2012), other, mostly cryptic species were newly described (e.g., Schlick-Steiner et al., 2003; Csősz et al., 2014; Seifert et al., 2014b; Seifert et al., 2017; Wagner et al., 2017) – and in some instances, cryptic diversity was made plausible without nomenclatural consequences (Seifert, 2009).

Harvesting ants of the genus *Messor* are ecosystem engineers, shaping their ecosystems by, among others, dispersing plant seeds, impacting seed banks, cycling nutrients, and modifying the microclimate (Plowes et al., 2013). For most of the 20th century, all Central European *Messor* populations were identified as *M. structor* (Latreille, 1798) (Fig. 1). Originally described from Brive-la-Gaillarde in France, *M. structor* was thought to occur in North Africa and the Middle East, Southern, Central, and Eastern Europe, Asia Minor, the Caucasus, and Central Asia (Czechowski et al., 2002; Lebas et al., 2016). However, morphological characters considered as diagnostic for *Messor structor*...
Fig. 1. Heuristics summary. (a) History. In 2006, the ants traditionally identified as *Messor structor* (Latreille, 1798) using qualitative morphology were suggested to comprise at least two mitochondrial DNA (mtDNA) lineages, A and B, based on a Central European sample (Schlick-Steiner et al., 2006). (b) Results of the current project. Analyzing the data from single disciplines separately, mostly in an unsupervised fashion, and then comparing the results from the various disciplines (Appendix S1), a wider geographic sample (Fig. 2) was analysed. mtDNA findings were interpreted as representing seven lineages, 1 to 7 (Fig. 3) when using the reciprocal-monophyly criterion of species delimitation. The other criterion applied to mtDNA, the transition from species-level to population-level branching for the bPTP results (see Section 3.1 for details), confirmed the general picture, but split Lineages 2, 6, and 7 in each two further lineages (Fig. S2). Traditional morphometrics (TM), using the phenotypic-distinctness criterion, came up with Lineages 1, 2, and 7, but grouped Lineages 3 to 6 together (Fig. 3); also, some individuals of Lineage 2 clustered with Lineages 3 to 6 and vice versa (indicated by coloured circles). Applying an endosymbiont-distinctness criterion to the Wolbachia data separated Lineages 2 and 6 but combined Lineages 3 to 5 (Fig. 3); no Wolbachia was found in Lineages 1 and 7. Genome scanning using amplified fragment length polymorphism (AFLP) separated Lineages 1, 2, 6, and 7, but grouped together Lineages 3 + 4 + 5, both when using a genotypic-clusters criterion for the BAPS results (Fig. 2, Fig. S5) and a reciprocal-monophyly criterion for the neighbor-net network results (Fig. 3) (see Section 3.4 for details). Applying an ecological-niche divergence criterion to the ENM data was the only supervised approach to species delimitation here, using the AFLP-based hypotheses. In the niche-identity tests, all these hypotheses were confirmed; in the background tests, Lineage 7 was significantly divergent from all other lineages (*). (c) Final species hypotheses. Considering all results, five species were delimited, congruent with Lineages 1, 2, 3 + 4 + 5, 6, and 7. (d) Reanalysis of data. Data from TM were reanalysed in a supervised mode using the final species hypotheses. Using discriminant analysis (DA), TM discriminated the species with a success rate of 68.1–100.0%; the corresponding values from leave-one-out cross validation (in parentheses) were 59.7–97.7%. (e) Nomenclatural consequences. Based on biogeographic arguments and type analyses using TM, Lineages 1, 3 + 4 + 5, and 6 were identified as *M. ibericus* Santschi, 1925, *M. structor*, and *M. muticus* (Nylander, 1849) stat.rev., respectively; Lineages 2 and 7 are here described as *M. ponticus* sp.n. and *M. mcarthuri* sp.n., respectively.

(e.g., width of metasternal process and scape length, Agosti and Collingwood, 1987) vary considerably within and across colonies, without geographic correlation readily discernible (Schlick-Steiner et al. 2006). Accordingly, a list of 19 available taxa currently regarded as junior synonyms or subspecies of *M. structor* (Bolton, 2017) testifies the views of earlier authors of there being additional diversity. In 2006, the one-species hypothesis was challenged by mitochondrial DNA (mtDNA) data suggesting the existence of at least two species, informally termed Lineages A and B (Schlick-Steiner et al., 2006), in line with life-history differences across Central Europe (Schlick-Steiner et al., 2005b). Within both Lineage A and B, substructure was discernible in the phylogenetic tree, but the phylogeny was based on just a single marker, and the geographical sampling was incomplete – it could therefore not be evaluated whether this substructure was due to further cryptic species. Here, we aimed at resolving the taxonomic situation of *Messor structor* using a multi-disciplinary approach and an increased geographic sample. We addressed the question whether the pattern seen in the mtDNA data can be ascribed to intraspecific geographic variation or whether the separate lineages represent additional, separate species. Upon completion of species delimitation, we aimed at drawing nomenclatural consequences, the important final step in taxonomic heuristics. Publishing nomenclatural consequences ensures that the results of the heuristic process last much longer than non-taxonomic publications do (Carstens et al., 2013; Schlick-Steiner et al., 2014), but such nomenclatural changes have started only recently to be published in high-visibility journals (e.g., Satler et al., 2013; Wachter et al., 2015) on a more regular basis (Steiner et al., 2015).

In more detail, we applied the protocol for integrative taxonomy of Schlick-Steiner et al. (2010): We aimed at generating data for all specimens under all disciplines and analysed the data from all disciplines independently, in an unsupervised (hypothesis-free) rather than supervised (hypothesis-driven) fashion as far as possible. We compared the results from disciplines and tried to find evolutionary explanations...
Fig. 2. Partial map of the Western Palearctic showing (a) the sampled Messor colonies (except each one nest from Kazakhstan and the Kyrgyz Republic, ID, UIBK 15123 and 15145; Table S1) as well as the corresponding results from (b) mitochondrial DNA (mtDNA; phylogenetic reconstruction using Bayesian and maximum likelihood, ML, algorithms), (c) traditional morphometrics (TM; hierarchical clustering using the Ward method of Agglomerative Nesting), (d) screening for Wolbachia endosymbions (phylogenetic reconstruction using ML), and (e) amplified fragment length polymorphism (AFLP, neighbor-joining topology, admixture analysis using BAPS). For details on the results shown in (b), (c), and (d), see Fig. 3, for those shown in (e), see Fig. S5.
for conflicting results of different disciplines. When this was not possible, we used data from an additional discipline which we expected to add information independent from that previously available and continued the iterative process until plausible explanations for all disagreements among disciplines had been identified (Appendix S1). After defining the final species hypotheses, data from disciplines which had resulted in deviating hypotheses when analysed without using prior hypotheses but which are useful in routine species identification, were reanalysed in a supervised fashion using the final species hypotheses.

We used five disciplines. (1) MtDNA has been one of the most widely used molecular tools in evolutionary and molecular-ecological research (Barrowclough and Zink, 2009), including in ants (Moreau, 2009), although a range of problems are known to potentially afflict working with it, such as issues related to introgression (Funk and Omland, 2003) and unknowingly amplifying pseudogenes (Song et al., 2008). To link the recent work to that of Schlück-Steiner et al. (2006), we nevertheless used mtDNA here and carefully screened the mtDNA results for problematic issues. (2) Morphological analysis facilitates linking the results from other disciplines to Linnean nomenclature, in that it can be used to analyse old but taxonomically relevant type specimens that often cannot be analysed by molecular means (Schlick-Steiner et al., 2007). Such linking is crucial to taxonomy and is usually done in a supervised mode. Traditional morphometrics (TM) has allowed distinguishing species via multivariate statistics for which exclusive character traits as used in qualitative morphology could not be found (Seifert, 2009). In the recent past, analytical tools have been developed that allow using TM data also in an unsupervised approach and thus independently from any other discipline (Ezard et al., 2010; Seifert et al., 2014a). Finally, TM is also relevant to routine identification once species delimitation has been completed. (3) The host-association patterns of symbionts have been increasingly used as taxonomic character (Gueguen et al., 2010; Gebiola et al., 2012; Van Steenberghe et al., 2015). Wolbachia are bacteria obligately endosymbiotic in arthropods and nematodes (Werren et al., 2008); ants are common Wolbachia hosts (Russell et al., 2017). Wolbachia can trigger reproductive isolation and increase speciation rate in their hosts (e.g., Shoemaker et al., 1999) but has, as far as we know, been included in just one species-delimitation study (Hernández-Roldán et al., 2016). A small-scale pilot study revealed infection of at least some Messor "structor" nests (data not shown), and we therefore screened the entire sample. In interpreting the results, we took into account that, like other symbionts, while usually inherited maternally, Wolbachia can be transmitted horizontally across host species (e.g., Schuler et al., 2013). (4) Nuclear DNA data are desirable in species delimitation as they can contain the potentially richest information on the evolutionary history of organisms. Ideally, markers with a lower substitution rate than mitochondrial DNA but sufficient sequence variation among species (Schlick-Steiner et al., 2010) are available (ant example: Ward and Summich, 2012). Testing established loci in Messor "structor" was not successful, in that they turned out to be either multicopy loci or were invariant (e.g., elongation factor 1-alpha, wingless, internal transcribed spacer; data not shown). We therefore used the genome-scanning method amplified fragment polymorphism (AFLP; Vos et al., 1995), which can be used in virtually any organism, without genomic resources at hand, and which continues being used in species delimitation (e.g., Lucentini et al., 2011; Reeves and Richards, 2011; Lima et al., 2012; Arthofer et al., 2013; Dejaco et al., 2016; Wagner et al., 2017). (5) The advent of ecological niche modelling (ENM) strongly enhanced the possibilities of using ecology in taxonomy by means of multivariate statistics, and several studies have since used the software MAXENT (Phillips et al., 2006) to distinguish closely related species by their predicted distribution (e.g., Rissler and Apodaca, 2007; Ross et al., 2010). Recognition of the potentially confounding effects of differences in the environment available to allopatric organisms on niche characterization triggered the availability of the background-similarity test to evaluate such effects (Warren et al., 2008). The background-similarity test has been embraced by the taxonomic community (e.g., Andújar et al., 2014; Rato et al., 2015) but has, to our knowledge, not been used in the species delimitation of ants to date.

2. Materials and methods

Ants keying out as Messor structor (Latreille, 1798) or its junior synonym M. muticus (Nylander, 1849) according to the identification key for workers of Agosti and Collingwood (1987) were sampled from Spain, France, Italy, Switzerland, Germany, Slovenia, Croatia, Austria, the Czech Republic, Hungary, Romania, Bulgaria, Greece, Turkey, Ukraine, the Russian Federation, Armenia, Kazakhstan, and the Kyrgyz Republic (Fig. 2, Table S1). Specimens were killed and stored frozen in 96% ethanol. From about 500 nests sampled between 1994 and 2007, 128 nests were selected as the core sample for the project; the criteria used in selecting these nests were DNA quality and a minimum number of 10 worker individuals. Of these nests, 32, 44, and 32, could not be used for AFLP, TM, and Wolbachia analysis, respectively, due to problems such as DNA degradation and PCR misamplifications for unknown reasons (AFLP, Wolbachia) and because more individuals than scheduled had been destroyed in the DNA extraction (TM). For a further 73 nests, the geographic coordinates were used in ENM, based on their species identification via TM and mtDNA sequencing (data not shown) after the data for the other disciplines had been generated.

Three Messor cf. wasmanni Krausse, 1910 nests from Italy, Croatia, and Greece were added as outgroup in the phylogenetic reconstructions using mtDNA and AFLP data. The nomenclature for this species followed Steiner et al. (2011) and not Schlück-Steiner et al. (2006), who treated the species under Messor concolor Santschi, 1927. Voucher specimens were deposited in the Hungarian Natural History Museum in Budapest (Hungary) and the University of Innsbruck (Austria). For type repositories of the newly described species, see Section 3.B.

2.1. TM

For traditional morphometrics of workers, 378 individuals belonging to 84 nests were analysed (for the procedure of selecting nests, see Section 2; Table S1). All measurements were made in µm using a pin-holding stage, permitting rotations around the X, Y, and Z axes. An Olympus SZX9 stereomicroscope was used at magnifications of 25× to 100× according to the size of the character, allowing a precision of ± 2 to 10 µm (S. Csösz, unpubl.). The following characters were recorded:

CL: Maximum cephalic length in median line excluding mandibles. The head must be carefully tilted to the position providing the true maximum.
CS: Absolute cephalic size; arithmetic mean of CL and CWb.
CWb: Maximum width of head capsule withoutcompound eyes. Measured posterior of the eyes.
FR: Frontal carinae distance. Distance between frontal carinae immediately caudal of posterior intersection points between frontal carinae and torular lamellae.
FS1: Maximum length of 1st funicular segment.
FS2: Maximum length of 2nd funicular segment.
FS3: Maximum length of 3rd funicular segment.
Och: Shortest diameter of oval compound eye.
Ocl: Longest diameter of oval compound eye.
ML: (Weber length): Mesosoma length from caudalmost point of propodeal lobe to transition point between anterior pronotal slope and anterior pronotal shield, measured in lateral view.
MW: Mesosoma width. Defined as longest width of pronotum in dorsal view.
NOH: Maximum height of petiolar node. Measured in lateral view from uppermost point of petiolar node perpendicular to a reference line extending from petiolar spiracle to imaginary midpoint of
transition between dorso-caudal slope and dorsal profile of caudal cylinder of petiole.

NOL: Length of petiolar node. Measured in lateral view from centre of petiolar spiracle to dorso-caudal corner of caudal cylinder. Do not erroneously take as reference point the dorso-caudal corner of the helicium, which is sometimes visible.

PEH: Maximum petiolar height. Chord of ventral petiolar profile at node level is reference line perpendicular to line describing maximum height of petiole.

PEL: Diagonal petiolar length in lateral view; measured from anterior corner of subpetiolar process to dorso-caudal corner of caudal cylinder.

PEW: Maximum width of petiole in dorsal view. Nodal spines are not considered.

PPH: Maximum height of postpetiole in lateral view. Measured perpendicularly to line defined by linear segment of border between dorsal and ventral petiolar sclerite.

PPL: Postpetiolar length. Longest anatomical line perpendicular to posterior margin of postpetiole and between posterior postpetiolar margin and anterior postpetiolar margin.


SL: Scape length. Maximum straight line scape length excluding articular condyle.

SW: Scape basal lobe width. Maximum width of the basal lobe of the scape.

TLD: Torular lamellae distance. Distance between distalmost edges of torular lamellae.

Species hypotheses were generated via the combined application of (i) Nest Centroid (NC) clustering and (ii) the Partitioning Algorithm based on Recursive Thresholding (PART). In this, the procedure applied in Csősz and Fisher (2016a, b) was followed; advantages and limitations of the procedure are discussed there. Briefly, (i) NC clustering uses discontinuities in continuous morphometric data to sort all similar cases into the same cluster in a two-step procedure. The first step reduces the dimensionality in the data via cumulative linear discriminant analysis (LDA) using nest samples (i.e., workers collected from the same nest are assumed genetically closely related) as groups (Seifert et al., 2014a). In the second step, pairwise distances between samples are calculated using as input the LD scores; the distance matrix is then displayed in a dendrogram. The NC clustering was done using cluster (Maechler et al., 2014) and MASS (Venables and Ripley, 2002). (ii) PART estimates the ideal number of clusters based on a recursive application of the Gap statistic (Tibshirani et al., 2001) and is able to discover both top-level clusters and sub-clusters nested within the top-level clusters. If more than one cluster is returned by the Gap statistic, it is reoptimized on each subset of cases corresponding to a cluster until a stopping threshold is reached or the subset under evaluation has less cases than twice the specified minimum cluster size (Nilsen et al., 2013). PART was computed with the package clusterGenomics v1.0 (Nilsen and Lingjaerde, 2013) using the function part, which also assigns observations (i.e. individuals or nests) into partitions. The clustering methods hclust and kmeans were used to determine the optimal number of clusters with 1000 bootstrap iterations using minSize set to 5 for hclust and 3 for kmeans. The results of PART were mapped on the dendrogram achieved under (i) by coloured bars via the function mark.dendrogram based on Beleites and Sergo (2015). The R script by Csősz and Fisher (2016a, 2016b) was used to implement (i) and (ii).

In the reanalysis of the TM data, the final species hypotheses were used in (confirmatory) LDA. In doing so, an exhaustive approach was used to identify the optimal character combination, that is the lowest number of characters resulting in the lowest classification error, following Moder et al. (2007). The result was validated using Leave-One-Out Cross-Validation (LOOCV). Classification hypotheses were imposed for all samples congruently classified by partitioning methods while wild-card settings (i.e. no prior hypothesis imposed on its classification) were given to samples that were incongruently classified by the two methods or proved to be outliers.

2.2. Mitochondrial DNA

DNA of single workers was extracted using the GenElute Mammalian Genomic DNA Miniprep Kit (Sigma, St. Louis, USA). The primers used for PCR amplification of the mitochondrial cytochrome c oxidase 1 (COI) gene were those used by Schlick-Steiner et al. (2006) and Steiner et al. (2011), except that the forward primer “Jerry” (Simon et al., 1994) was additionally used to amplify shorter but overlapping stretches of degraded samples. PCR conditions followed Steiner et al. (2011) except for a few degraded samples, for which 1.25 U of Restorease DNA Polymerase (Sigma) was incubated with 2.5 μl DNA extract, 1 × reaction buffer (provided by Sigma and including MgCl₂), 0.2 mM dNTPs and 18 μl dH₂O for 45 min at 37 °C. Then, 3.6 μM of forward and reverse primer was added, yielding a total reaction volume of 25 μl per sample, and the thermocycler program continued as follows: 3 min at 95 °C, 3 min at 60 °C, 35 cycles of 30 s at 95 °C, 45 s at 48 °C, and 2 min at 72 °C, and a final incubation of 10 min at 72 °C. Product purification and Sanger sequencing using the PCR primers followed either Steiner et al. (2011) or Wachtler et al. (2015). Following the recommendations of Song et al. (2008), all gels, electropherograms, and sequences were screened for indications of nuclear pseudogenes of mitochondrial origin (numts), that is, ghost bands, ambiguous peaks, frameshifts, premature stop codons, and compositional abnormalities. All mutation singletons were confirmed by independent PCR.

Eleven sequences from Schlick-Steiner et al. (2006) (DQ074330, DQ074332, DQ074334, DQ074339-DQ074341, DQ074350, DQ074353-DQ074355) were added, and a total of 208 sequences were aligned with CLUSTALX v2.0 (Larkin et al., 2007). Ninety-seven of the sequences had a length of 1375 bp and were collapsed to 58 haplotypes and used for the final mtDNA analyses. The other 111 sequences were of 850 bp length and were added to the alignment at a later point in the heuristic process (Appendix S1). This was done to identify their mtDNA lineage via additional phylogenetic reconstructions using a maximum likelihood (ML) algorithm. The geographic coordinates of those 111 sequences were then used in ENM.

For phylogenetic reconstructions using ML, model selection was performed with MODELTEST v0.1.1 (Posada, 2009). Using the cumulative Akaike and the Bayesian information criteria resulted in the TIM3 + G + I and the HKY + I model as best fit for the data, respectively. Both models were used in the ML-based reconstructions as implemented in MEGA v6.0.6 (Tamura et al., 2013) using the settings subtree-pruning-regrafting = extensive (SPR level 5), make initial tree = automatically (NJ), branch swap filter = very strong. In evaluating node support, a significance threshold of > 75% was applied (Soltis and Soltis, 2003). MEGA cannot perform separate parameter estimations for partitions according to codon position. We thus additionally constructed ML trees using the web interface of IQ-TREE (Trifinopoulos et al., 2016) for this purpose. Both the TIM3 + G + I and the HKY + I model were tested in multiple runs under various “perturbation strength” and “IQ-tree stopping rules” settings. As the support for the single lineages was corroborated in most runs, the detailed results are not shown. Additional trees were computed as input for a Bayesian implementation of the Poisson tree processes (see below).

Bayesian Inference (BI) was performed with MrBayes v3.2.2 (Ronquist et al., 2012) using the mixed model and partitions according to codon position. Two parallel runs with four metropolis coupled Markov chains each, with temperature default settings, were sampled every 100 generations for 4,000,000 generations. After 3,000,000 generations, the standard deviation of split frequencies remained below 0.005; thus, the last 10,000 trees were used for consensus tree construction. The convergence check with TRACER v1.6 (http://beast.bio.ed.ac.uk/Tracer) supported this decision. A significance threshold for
node support of posterior probability > 0.95 was used (Huelsenbeck and Rannala, 2004). Genetic variation within and between lineages was computed with MEGA v6.0.6 based on Tamura-Nei distance.

Following Pons et al. (2006) and Fontaneto et al. (2007), the Generalized Mixed Yule-Coalescent (GMYC) approach for delimiting species was applied. In doing so, an ultrametric tree based on the HKY + I model (using partitions according to codon position) was constructed in BEAST v1.8.0 (Drummond et al., 2012), sampling every 100 generations for 10,000,000 generations. Convergence was checked with TRACER v1.6. The first 2,500,000 generations were discarded as burnin, and the remaining trees were summarised via TREEANNOTATOR v1.8 (http://beast.bio.ed.ac.uk/treannotator). The R package SPLITS (Ezard et al., 2009) in R v3.0.2 (R Development Core Team, 2011) was used to carry out single- and multiple-threshold analyses (Monaghan et al., 2009).

The Bayesian implementation of the Poisson tree processes (bPTP) model of Zhang et al. (2013) was reported to be less error-prone than the GMYC approach (Zhang et al., 2013) and was therefore additionally conducted. The analyses were performed via the web server at http://species.h-its.org/ptp/ using as input the ML tree. Because the number of haplotypes across supported clades was uneven in this tree and because bPTP may fail on unbalanced datasets (Zhang et al., 2013), multiple tests were run with trees after thinning haplotypes. This did not influence the bPTP output, and the complete ML tree was used in the final run.

### 2.3. AFLP

AFLP data were generated following the protocol of Vos et al. (1995) with modifications, using the DNA extracts as specified under mtDNA. DNA digestion and adaptor ligation were performed in a mixture containing 1 × T4 Ligase Buffer (all AFLP enzymes by Thermo Scientific, Waltham, USA), 1 µl 0.5 M NaCl, 0.5 µl 0.1% BSA, 1 U MseI, 5 U EcoRI, 1 U T4 DNA Ligase, each 1 µl conditioned MseI (50 µM) and EcoRI (5 µM) adapters and approx. 50 ng template DNA in a total volume of 10 µl. Reactions were incubated at 37 °C for 2 h and then diluted 1:10 with TLE (20 mM Tris, 0.1 mM EDTA, pH 8.0). Pre-selective amplicons were generated using 1 × MyTaq Buffer (Bioline, London, UK), 0.2 µM of both Mse-C and Eco primer, and of 0.25 U MyTaq polymerase in 10 µl total volume. Amplification conditions for the preselective PCR were 72 °C for 2 min, 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 2 min, and a final elongation at 60 °C for 10 min. The product was diluted 1:10 with TLE, and selective PCR was carried out with the primer pairs Mse-CTA & Eco-AC, Mse-CAA & Eco-AA, and Mse-CTT & Eco-AA. To introduce FAM/PET/NED fluorophores to the PCR products, we did M13-tailing as described in Wachter et al. (2012). Success of both amplification steps was verified by agarose gel electrophoresis.

Fragment analysis of selective amplicons was performed on an ABI 3130 capillary sequencer (Applied Biosystems, Foster City, USA) by a commercial provider. PeakScanner v1.0 (Applied Biosystems) was used for electrophrogram visualisation and data export. optiFLP v1.54 (Arthofer et al., 2011) was used to identify preliminary scoring parameters (Table S2), and all samples with a peak frequency score in the 10th percentile in more than one primerset as calculated with tinyFLP v1.40 (Arthofer, 2010) were excluded. optiFLP was re-run on this reduced set, giving the final scoring parameters (Table S2). tinyCAT v1.20 (Arthofer, 2010) was used to concatenate the three binary matrices and generate input files for the subsequent analyses.

For Bayesian clustering, both BAPS v5.3 (Corander et al., 2008) and STRUCTURE v2.3.3 (Pritchard et al., 2000) were used. In BAPS, 10 repetitions for each K = {1, 10} were performed, and K = 8 was identified as optimum partition. The topology of the clusters was visualised with the neighbor-joining algorithm implemented in BAPS using Nei distances. The results of the cluster analysis served as input for BAPS admixture analysis based on 500 simulations from posterior allele frequencies.

In STRUCTURE, 10 runs for each K = {1, 10} were performed with 200,000 steps, discarding the first 20,000 steps as burnin. To foster convergence, the mtDNA clade of each sample was provided as locprior. The best K was determined using the algorithm described by Evanno et al. (2005) as implemented in STRUCTURE HARVESTER v0.6.94 (Earl and Vonholdt, 2012).

A neighbor-net network was constructed in SplitsTree v4.13.1 (Huson and Bryant, 2006) using the settings character transformation = Jaccard distances, distance transformation = neighborNet with ordinary least squares, splits transformation = EqualAngle using weights and Convex Hull option. In the same program, a neighbor-joining tree was calculated using the settings character transformation = Jaccard distances, distance transformation = NJ, splits transformation = Phylogram. The tree was exported as Newick file, and FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/) was used to collapse and rotate nodes.

Reconstructing a Bayesian tree from the concatenated AFLP data was attempted using MrBayes v3.2.2 with the following settings: coding = noobsencesites, statefreq = Dirichlet with the 0:1 ratio in the binary AFLP matrix, Nruns = 2, Nchains = 6, temp = 0.01. After 120,000,000 generations, no convergence had been approached, and the analysis was terminated (data not shown).

### 2.4. Wolbachia endosymbionts

For Wolbachia diagnosis, the primers wsp151F and wsp691R (Ruang-arereate and Kittayapong, 2006) were used in 5 µl PCR reactions containing 1 × MyTaq reaction buffer, 0.2 µM of each primer, and 0.125 U MyTa polymerase. The cycling scheme was 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min, and a final elongation at 60 °C for 10 min. All negative reactions were repeated using the QIAGEN (Hilden, Germany) Probe PCR mastermix with the same primers and cycling conditions, and samples without band in either PCR were considered as uninfected. Multilocus sequence typing (MLST) of infected samples followed the standard protocols (Baldo et al., 2006) with modified annealing temperatures (ftz, gatB: 47 °C; fipA: 48 °C; coxA: 50 °C; hcpA: 55 °C) and using MyTaq PCR chemistry. Visual inspection of the sequence electropherograms using Chromas Lite v2.1.1 (Technelysium Ltd., South Brisbane, Australia) gave no hints of multiple infections. Thus, the direct MLST and wsp sequences of each strain were concatenated, resulting in 2666 bp, and aligned with ClustalX v2.0. Model selection and ML tree reconstruction with the HKY model were performed in MEGA v6.0.6. Node support was evaluated using 1000 bootstrap pseudo-replicates and a significance threshold of > 75%.

### 2.5. ENM and ENM-based tests for niche divergence

The interplay between “Eltonian” processes (biotic interactions), “Grinnellian” factors (broad-scale abiotic environmental conditions), and dispersal or migration defines the realised niche (Soberon, 2007). Here, climate, soil, and vegetation variables were considered a reasonable approximation of the realised niche. Nineteen bioclimatic data layers (bio1–19) for current conditions (1960–2000) were compiled from the WorldClim database v1.4 (http://www.worldclim.org; Hijmans et al., 2005). Eight soil variables (https://soilgrids.org; Hengl et al., 2014, 2017) and five remote-sensing-derived vegetation landcover products (Tuanmu and Jetz, 2014) were selected due to their importance for the distribution of ant species (Johnson, 1992, 2000; Boulton et al., 2005; Rios-Casanova et al., 2006; Dattilo et al., 2013; Wang et al., 2016). Also, the inclusion of remote-sensing-derived vegetation data potentially provides information about Eltonian niche characteristics in the broadest sense (McCormack et al., 2010; Tuanmu and Jetz, 2014). As a consequence, ENM was expected to capture more detailed aspects of the realised niches of Messor lineages than climate
alone.

All data layers were compiled at 30-arc-second resolution and clipped to the area encompassing the distribution of *Messor* "*structor*" (Czechowski et al., 2002; Lebas et al., 2016). In doing so, the poorly sampled Central Asian regions east of the Caspian Sea (and thus the localities of Lineage 6 in Kyrgyzstan and Kazakhstan; Table S1) were excluded from ENM to reduce sampling bias and background selection issues (Phillips et al., 2009; Barbet-Massin et al., 2012; Merow et al., 2013). ENMTools v1.4.4 (Warren et al., 2010) was used to assess pairwise correlation between variables. Highly correlated variables (Pearson's correlation coefficient > 0.95) were removed; 28 variables were retained (Table S3).

ENM was the only discipline that needed prior hypotheses on taxon identities; the AFLP-based hypothesis of five lineages was used, that is, Lineages 1, 2, 3 + 4 + 5, 6, and 7 (for details, see Section 3). From about 500 nests sampled, only those were considered for ENM for which molecular and morphological data were available. Duplicate presence records and localities within a spherical distance of 5 km were removed to avoid model overfitting due to spatial autocorrelation (Veloz, 2009; Hijmans, 2012; Boria et al., 2014) using PASSaGE2 v2.0.11.6 (Rosenberg and Anderson, 2011). The final dataset consisted of 110 records, including 28 for Lineage 1, 15 for Lineage 2, 33 for Lineages 3 + 4 + 5, 22 for Lineage 6, and 12 for Lineage 7. The niche space between the five lineages and 1000 background points was compared using principal component analysis (PCA) in R (rcomp command). The PCA was performed on the matrix of 28 variables, and the component scores of each data point were projected in two dimensions. Based on the selected 28 variables, ecological niche models were constructed for each lineage using default parameters in MAXENT v3.3.3k (Phillips et al., 2006).

Niche identity tests and background similarity tests, hereafter identity and background tests, were designed to complement each other and to reflect the continuum represented by niche similarity (Warren et al., 2008, 2010; McCormack et al., 2010; Martínez-Cabrera et al., 2012). As a consequence, both tests were used to assess whether lineages have diverged ecologically (Warren et al., 2008; McCormack et al., 2010). The identity test examined the null hypothesis that each pair of taxa (here pairs among the five lineages) is distributed in identical environmental space by comparing similarities of the actual niche models generated with actual occurrence localities of two species to pseudo-niches generated with points randomly selected from a pool of occurrence localities of the same taxa (Warren et al., 2008, 2010). Niches were considered significantly different if the observed value of niche overlap was below the 95% confidence limits of the null distribution of niche overlap values generated from the random pseudo-replicates (Warren et al., 2008). A rejection of niche identity, however, may be driven by differences in environmental conditions within the available habitat of two taxa.

The background test compares differences in the environmental background to determine if two taxa are more or less similar than expected based on their environmental background (Warren et al., 2008; McCormack et al., 2010). For each combination of lineages, the niche model for the focal lineage was compared with a series of pseudo-replicate models generated by randomly sampling the background of the other lineage and vice versa. The null hypothesis that the degree of niche overlap of two lineages is explained by the available habitat is rejected if the observed value of niche overlap is lower (=niche divergence) or higher (=niche conservatism) than the 95% confidence limits of the null distribution (Warren et al., 2008; McCormack et al., 2010). Failure to reject the null hypothesis, however, does not imply that there is no niche divergence or conservatism but rather that it cannot be disentangled whether niche differences are biologically meaningful or simply arise from geography alone (Warren et al., 2008; Nakazato et al., 2010; McCormack and Maley, 2015). For each lineage, a biologically realistic background was defined as circular buffer of 300 km in diameter around the known occurrence points (Warren et al., 2010).

The identity and background tests were performed in ENMTools with 200 replicates each. Pairwise niche similarity was assessed using Schoener's D, which was found to outperform Hellinger's J (Rödler and Engler, 2011). The null distributions and associated confidence limits were visualised using the profiles.plot function of the R package diversitree v0.9-8 (FitzJohn, 2012).

### 2.6. Species concept and species delimitation criteria

The unified species concept (de Queiroz, 2007) was applied, under which a separately evolving metapopulation lineage is the only necessary conceptual property of species. The species delimitation criteria used were: mtDNA – reciprocal monophyly (cf. Donoghue, 1985) for the phylogenetic trees, the transition from species-level to population-level branching for the bPTP and GMYC results; TM – phenotypic distinctness (cf. Sokal and Crovello, 1970); Wolbachia – endosymbiotic distinctness; AFLP – reciprocal monophyly for the phylogenetic trees and neighbor-net networks, genotypic clusters (Mallet, 1995) for the STRUCTURE and BAPS results; ENM – ecological niche divergence (cf. Vanvalen, 1976; cf. Funk et al., 2006).

### 3. Results and discussion

The integrative-taxonomic workflow consisted of three iterations, starting with mtDNA, TM, and Wolbachia, then including also AFLP and finally ENM (see Appendix S1 for details). Here, the heuristic process is neglected, and the results from the various disciplines are summarized (Fig. 1).

#### 3.1. MtDNA

No indications of numts were detected. The Bayesian and ML-based phylogenetic reconstructions using mtDNA sequences revealed seven well supported clades, Lineages 1–7 (Fig. 3a), of which Lineages 1–2 and Lineages 3–6 reflected the substructures within Lineages A and B of Schlick-Steiner et al. (2005b), respectively. Lineage 7 was newly identified from the south-eastern region of the study area, a region not analysed in the 2005 study. Variation within lineages ranged from 0.2% (Lineages 1, 3, 4) to 1.8% (Lineage 7), variation between lineages from 2.5% (Lineages 1–2) to 9.2% (Lineages 2–4).

GMYC (Fig. S1) and bPTP (Fig. S2) resulted in more entities at ten and nine lineages, respectively, than the BI and ML approaches. While GMYC and bPTP are welcome as objective means of delimiting species from single-locus gene trees, they are also known to be influenced by sampling density and effective population size and seem to be prone to over-splitting (Satler et al., 2013; Ahrens et al., 2016). We consider this as plausible also here, that is, we further use only the seven-lineages hypothesis among the mtDNA results.

#### 3.2. TM

NC clustering combined with two clustering methods (kmeans and hclust) of PART using the TM data generated a clear-cut four-cluster hypothesis (Fig. S3). Samples of Lineages 1 and 7 formed distinctive clades each. Lineage 2 formed a cluster of its own, but two samples belonging to Lineage 6 were nested in this subset. Samples of Lineages 3–6 grouped together and intermingled with samples of Lineage 2 in some instances (Fig. 3b).

#### 3.3. Wolbachia

Wolbachia infections were found in all lineages except Lineages 1 and 7 (Fig. 2). The infection rates of lineages ranged from 30 to 100% (Table S1). MLST sequencing revealed four different strains (Fig. 3c). The phylogeographically most distinct strain (2.95% sequence divergence
from its closest relative in the concatenated MLST & wsp alignment) was exclusive to Lineage 6 and was thus termed wMes6. Another strain was exclusive to Lineage 2 (wMes2), and the remaining two strains wMes3 and wMes4 occurred in multiple lineages with highest abundance in Lineages 3 and 4, respectively (Fig. 3c).

3.4. AFLP

The neighbor-joining tree based on 513 AFLP loci revealed five partly supported clades, reflecting Lineages 1, 2, 6, and 7, and another clade merging Lineages 3 + 4 + 5 (Fig. 3d). As the clades are separated by quite some distance, a more sophisticated tree construction might have yielded significant support to all clades, but multiple MrBayes runs following the recommendations of Koopman et al. (2008) did not converge (standard deviations of split frequencies not below 0.047 after 120 million generations), emphasizing the need for improved methods to reconstruct phylogenies from AFLP data (Graves, 2009; Mrinalini et al., 2015). The five clades and the intermingling of samples from Lineages 3 + 4 + 5 in one of the clades were also visible in the neighbor-net network (Fig. 3d), which is, despite lacking a measure of significance, a helpful tool for data exploration (Bryant and Moulton, 2004). Genotypic cluster analyses via STRUCTURE and BAPS yielded contradicting outcomes – while the first approach divided the data into just two clusters (Fig. S4), the latter split it into 13 clusters (Fig. 2, Fig. S5). However, the problems of STRUCTURE in identifying the real number of clusters are well known, especially if sample sizes are uneven (Kalinowski, 2011; Falush et al., 2016; Puechmaille, 2016). BAPS, in contrast, is less sensitive to sample-size variation and tends to detect finer genetic differentiation (Wilkinson et al., 2011), thus potentially oversplitting in species delimitation (e.g. Dejaco et al., 2016). In accounting for this, we merged the BAPS-clusters according to their position within monophyletic clades in the BAPS tree. This resulted in five clades identical to those of the phylogenetic reconstruction using mtDNA. We concluded that the five clades were the most plausible
Fig. 4. Results from ecological niche modelling. (a) Principal component analysis biplot of the occurrence data of the *Messor* lineages (coloured squares) and 1000 background points (grey) drawn at random from the distribution of the *Messor* "*structor*" group and based on 28 climate, soil, and vegetation variables (Table S3). The first two principal components (PC1 and PC2) accounted for 52% of the total variance. (b) Potential distribution maps of *Messor* "*structor*" lineages obtained with MAXENT, based on their known distribution (black crosses) and climate, soil, and vegetation variables. Colours indicate probabilities of suitability above the maximum training sensitivity plus specificity logistic threshold. (c) Background tests of niche divergence. Observed niche overlap values of Schoener’s *D* for each pairwise combination of lineages (black dashed line) are compared with null distributions of background divergence. C, significant niche conservatism (*P* < 0.05); D, significant niche divergence (*P* < 0.05).
result of the AFLP approach.

3.5. ENM and niche divergence

In all ecological-niche analyses, the AFLP-based five-lineages hypothesis (Lineages 1, 2, 3 + 4 + 5, 6, and 7) was used. In the PCA (Fig. 4a), no single variable contributed strongly to the loadings of the first two principal components, but PC1 roughly corresponded to soil and temperature and PC2 to seasonality and precipitation variables (Table S4). As a tendency, Lineages 1, 2, 3 + 4 + 5, 6, and 7 were separated, but their clusters were abutting and even overlapping.

The ecological-niche models performed well for all five lineages with Area Under the Curve values of 0.975 for Lineage 1, 0.989 for Lineage 2, 0.979 for Lineages 3 + 4 + 5, 0.989 for Lineage 6, and 0.985 for Lineage 7. ENM suggested unique geographic distributions for all lineages (Fig. 4b). Niche overlap was relatively low (Schoener’s D: 0.02–0.41; Fig. 4c), which was confirmed by the identity tests that returned significant ecological differentiation between all lineages (Fig. S6). The background test confirmed this result for most pairs of Messor lineages. In more detail, for the lineage pairs 3 + 4 + 5 vs. 7, and 6 vs. 7, this divergence was significant in both directions; significant divergence in one direction emerged between Lineages 1 vs. 3 + 4 + 5, 1 vs. 7, 2 vs. 7, and 3 + 4 + 5 vs. 6 (Fig. 4c) – significance in one direction is already considered as stringent result, however (McCormack et al., 2010; Andújar et al., 2014). For the remaining lineage pairs (1 vs. 2, 1 vs. 6, 2 vs. 3 + 4 + 5, 2 vs. 6), the observed niche overlap was not significantly different from the null distributions based on the available habitat (Fig. 4c).

ENM approaches that account for environmental divergence of spatially auto-correlated background habitats can detect adaptation to different ecological niches between candidate species and thus provide an important tool for species delimitation in integrative taxonomy (McCormack et al., 2010; Nakazato et al., 2010; Andújar et al., 2014; Lopez-Alvarez et al., 2015; Rato et al., 2015). Here, the niche identity tests revealed a clear differentiation between the realised niches of all five lineages (Fig. S6). For species with at least partially differing distributions, a certain degree of niche differences may be inevitable given the regional differences in environmental niche variables (Warren et al., 2008; McCormack et al., 2010; Nakazato et al., 2010). However, in the present case, the divergence was significantly stronger than expected based on the available background habitat for most (six out of ten) Messor lineage combinations (Fig. 4c). It cannot be disentangled for the four exceptions (lineage pairs 1 vs. 2, 1 vs. 6, 2 vs. 3 + 4 + 5, 2 vs. 6) whether the significant niche difference observed in the identity test (Fig. S6) is biologically meaningful or within the bounds expected due to regional differences in niche variables (McCormack et al., 2010; Nakazato et al., 2010; McCormack and Maley, 2015).

3.6. Combining evidence, explaining disagreements among disciplines, and generating final species hypotheses

By combining the evidence from the various disciplines used, the question raised at the outset can be answered whether the substructure of Lineages A and B as characterised by Schlick-Steiner et al. (2006) represents separate species or intraspecific variation. When applying the species-delimitation criteria we defined for the various disciplines, the substructure within Lineages A (Lineages 1, 2) and B (Lineages 3, 4, 5, 6) seen in the mtDNA data does, at least partly, represent additional, separate species as does the newly identified Lineage 7, outside Lineages A and B. Deciding where exactly the new species limits should be drawn is more difficult, though, given the disagreements between some of the disciplines. Generally, such disagreements arise when properties of organisms used to delimit species evolve at different time points in the speciation process, an effect well accommodated by the unified species concept (e.g., Fig. 1 in de Queiroz, 2007). In the following, the seven lineages are discussed in the order of increasing
For Lineages 3 + 4 + 5, the strongest disagreements among disciplines were returned (Fig. 1): In mtDNA, they emerged as three entities separate from each other and from all other lineages, both when applying a reciprocal-monophony and a species-population-transition criterion. In TM, they were not separable from each other and from Lineage 6 and just poorly from Lineage 2 (but see under Lineages 2 and 6 that this may not be an important finding for cryptic species). In Wolbachia and AFLP, no separation at all was found between Lineages 3, 4, and 5, but against all other lineages, except for one individual of Lineage 6 infected with wMes3. Finally, in ENM, Lineages 3 + 4 + 5 were clearly separable from all other lineages with the exception of Lineage 2 (see the discussion under Lineage 2). We argue that, among the molecular-genetic data, the clear-cut AFLP finding of a single separate entity should be prioritised over the clear-cut mtDNA finding of three separate entities: Peripatric speciation of Lineage 6 out of the easternmost distribution range of the ancestor of Lineages 3 + 4 + 5 would be geographically well imaginable and would be expected to result in paraphyly of Lineages 3 + 4 + 5 relative to Lineage 6 (Funk and Omland, 2003). We conclude that Lineages 3 + 4 + 5 represent one species separate from all other lineages analysed here, with the highest within-species mtDNA variation of all species at 2.3%.

Under the unified species concept (de Queiroz, 2007) and the species-delimination criteria we applied (for details, see Section 2.6), the geographic distribution of lineages is not relevant. However, under the biological species concept (Mayr, 2000), sympatry of distinct lineages is interpreted as a hint of reproductive isolation (debated by others for theoretical and empirical reasons; for a brief summary, see, e.g., Steiner et al., 2010 and references therein). The coarse geographic sampling design of the current study does not allow a consistent evaluation of sympathy (Fig. 2a). However, of the 10 pairwise combinations of the five species, two were found co-occurring on the spot (Lineages 2 and 3 + 4 + 5; 2 and 6) and three in distances between 22 and 35 km, making sympathy very plausible (Lineages 1 and 2; 1 and 3 + 4 + 5; 2 and 7). Four pairwise combinations were found in distances between 105 and 274 km (Lineages 1 and 6; 1 and 7; 3 + 4 + 5 and 6; 3 + 4 + 5 and 7). Finally, the shortest distance between records of Lineages 6 and 7 was 592 km, and these two species may indeed not occur in sympatry.

3.7. Qualitative-morphological analysis and reanalysis of TM data

The decision on Lineages 1, 2, 3 + 4 + 5, 6, and 7 to represent separate species facilitated searching for species-specific qualitative-morphological characters. In this supervised analysis of workers and queens, differences were found to discriminate all species, except Lineages 3 + 4 + 5 from 6, based on body colour, base-of-scape’s shape, surface characteristics of anepisternum, mesopleuron, and first gastral tergite, and head setation (Fig. 5; for details, see Section 3.8).

Availability of the final species hypotheses also facilitated reanalysis of the TM data in a supervised fashion, which can allow separating species indistinguishable in unsupervised analyses (Schlick-Steiner et al., 2010). In the current case, the highest success rate achieved in LDA, using 15 characters (FR, FS1, FS3, HL, MW, NOH, OcH, Ocl, PEW, PPH, PPL, PPW, SL, SW, TLD; Table 1) identified by the optimal-distinction approach (Modé et al., 2007), was 87.6% of worker individuals over all five species (from 68.1% to 100.0% per species; Fig. 1), and LOOCV returned 84.1% overall error rate (59.7% to 97.7% per species; Fig. 1). Importantly, morphologically discriminating individuals of Lineages 6 and 7 became feasible; using the discriminant formula of Root 1 perfectly separates the two species (see Table 1 for details on the calculation). However, Lineages 3 + 4 + 5 and Lineage 6 still can be discriminated just poorly and, overall, the error rates achieved are higher than for various other groups of cryptic ant species using comparable character systems (Steiner et al., 2010; Csőz et al., 2014; Seifert et al., 2014b, 2017; Csőz and Fisher, 2016a; Seifert, 2016; Wagner et al., 2017). However, the further we advance in decrypting cryptic diversity, the more frequently we will encounter species that cannot be distinguished morphologically, including in supervised analyses (e.g., Dejeaco et al., 2016). Use of hitherto untested morphological characters and/or ways of assessing characters may open up new possibilities in the future (cf. Schlick-Steiner et al., 2010). For example, analyzing male morphology, and especially male genital morphology (ant example: Wagner et al., 2017), and/or quantitatively analyzing shape as possible under the framework of geometric morphometrics (ant example: Bagherian Yazdi et al., 2012) may improve species discrimination. Finally, analyzing nuclear DNA in a more sophisticated fashion, ideally via whole-genome sequencing, with an increasing number of ant genomes available as background information (Boomsma et al., 2017), will make possible testing the species hypotheses presented here.

3.8. Nomenclatural consequences and species descriptions

Based on the geographic position of the type locality of Messor structor (Latreille, 1798) (Brive-la-Gaillarde, France) and the geographic distributions of Lineages 1 to 7 established here (Fig. 2), M. structor was identified as name for Lineage 3 + 4 + 5 (for details, see Appendix S2). A sample from Brive-la-Gaillarde was designated as neotype under the terms of Art. 75 of the International Commission on Zoological Nomenclature (1999) (for details, see Section 3.8.3).

In identifying names for the remaining four species delimited here,
Table 1
Reanalysis of traditional-morphometric data achieved by Linear Discriminant Analysis using 15 characters run on *Messor ibericus*, *M. ponticus* sp.n, *M. structor*, and *M. mcarthuri* sp.n. (a) Unstandardized row coefficients of Roots 1–4. For definitions of characters, see Section 2. (b) Geometric means of discriminant scores for nest samples calculated from individuals of each species achieved by Roots 1–4. To use the values for species identification, each morphometric value [µm] from the characters in the first column of (a) has to be multiplied by the corresponding value in the column of the root of interest. Then, the sum of all 15 products and the constant of the root column is calculated. The result can be compared with the values in (b) to determine the species identity of the individual.

<table>
<thead>
<tr>
<th>Root 1</th>
<th>Root 2</th>
<th>Root 3</th>
<th>Root 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) CL</td>
<td>−0.0039</td>
<td>0.0033</td>
<td>0.0012</td>
</tr>
<tr>
<td>FR</td>
<td>0.0162</td>
<td>0.0250</td>
<td>−0.0246</td>
</tr>
<tr>
<td>FS1</td>
<td>0.0488</td>
<td>0.0605</td>
<td>0.0024</td>
</tr>
<tr>
<td>FS3</td>
<td>−0.0303</td>
<td>0.0114</td>
<td>−0.0428</td>
</tr>
<tr>
<td>MW</td>
<td>−0.0105</td>
<td>0.0019</td>
<td>−0.0244</td>
</tr>
<tr>
<td>NOH</td>
<td>0.0171</td>
<td>−0.0209</td>
<td>0.0036</td>
</tr>
<tr>
<td>OCH</td>
<td>0.0104</td>
<td>−0.0285</td>
<td>0.0095</td>
</tr>
<tr>
<td>OCL</td>
<td>0.0107</td>
<td>−0.0028</td>
<td>0.0095</td>
</tr>
<tr>
<td>PEW</td>
<td>−0.0078</td>
<td>0.0172</td>
<td>0.0018</td>
</tr>
<tr>
<td>PPH</td>
<td>−0.0043</td>
<td>0.0298</td>
<td>−0.0141</td>
</tr>
<tr>
<td>PPL</td>
<td>0.0177</td>
<td>−0.0074</td>
<td>−0.0049</td>
</tr>
<tr>
<td>PPSW</td>
<td>0.0045</td>
<td>−0.0119</td>
<td>0.0268</td>
</tr>
<tr>
<td>SL</td>
<td>−0.0136</td>
<td>−0.0029</td>
<td>0.0149</td>
</tr>
<tr>
<td>SW</td>
<td>0.0677</td>
<td>−0.0558</td>
<td>−0.0151</td>
</tr>
<tr>
<td>TLD</td>
<td>−0.0055</td>
<td>0.0398</td>
<td>0.0563</td>
</tr>
<tr>
<td>Constant</td>
<td>−4.1538</td>
<td>3.2448</td>
<td>−2.6209</td>
</tr>
</tbody>
</table>

(b) M. ibericus

<table>
<thead>
<tr>
<th>M. ibericus</th>
<th>0.0162 ± 0.036 [−2.635, −1.647]</th>
<th>−1.204 ± 0.48 [−2.184, −0.673]</th>
<th>1.575 ± 0.58 [0.710, 2.337]</th>
<th>−0.763 ± 0.42 [−1.510, −0.280]</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. ponticus</td>
<td>80.821 ± 0.92 [−1.868, 0.572]</td>
<td>−2.179 ± 1.02 [−3.527, −0.655]</td>
<td>−1.621 ± 0.52 [−2.342, −0.923]</td>
<td>0.016 ± 0.99 [−2.455, 1.147]</td>
</tr>
<tr>
<td>M. structor</td>
<td>0.922 ± 0.52 [−0.313, 1.865]</td>
<td>0.136 ± 0.52 [−1.057, 0.923]</td>
<td>0.330 ± 0.60 [−0.734, 1.683]</td>
<td>0.542 ± 0.72 [−0.886, 1.861]</td>
</tr>
<tr>
<td>M. mcarthuri</td>
<td>1.813 ± 0.77 [0.722, 3.181]</td>
<td>0.649 ± 0.92 [−2.084, 1.498]</td>
<td>−0.532 ± 1.30 [−3.965, 0.769]</td>
<td>−1.053 ± 0.70 [−2.427, 0.327]</td>
</tr>
<tr>
<td>M. mcarthuri muticus</td>
<td>4.062 ± 0.90 [−5.379, −2.660]</td>
<td>1.552 ± 0.92 [−1.302, 3.212]</td>
<td>−0.667 ± 0.48 [−1.266, 0.278]</td>
<td>0.013 ± 0.63 [−1.190, 1.142]</td>
</tr>
</tbody>
</table>

3.8.1. Diagnosis. *Worker and queen.* Colour generally lighter than in *M. muticus* and *M. ponticus* sp.n. Rugosity of cuticular surface more regular than in *M. structor*, rugae sometimes parallel. Discrimination from *M. structor* and *M. muticus* by short 1st funicular segment and by base of scape without lobe but with a small tooth-like processes. Discrimination from all four species by finely costate base of mesepimeron. Head costate almost throughout unlike in *M. ponticus*, and costae less regular than in *M. mcarthuri* sp.n. Setae more abundant on side of head than in *M. muticus*, similarly to *M. structor*, *M. muticus*, and *M. mcarthuri*. Petiole not as costulate as in *M. mcarthuri*. Discrimination from *M. ponticus*, *M. structor*, and *M. muticus* by 1st gastral tergite’s entirely imbricate surface.

*Worker.* Size similar to *M. structor* and *M. muticus*, major workers usually smaller than in *M. ponticus* sp.n. and *M. mcarthuri* sp.n. For individuals difficult to discriminate from *M. structor* using qualitative morphology, linear discriminant function D2 ibericus vs. structor is available, which uses two morphometric characters with a classification success of 96.8%:

\[
\text{D2 ibericus vs. structor} = 0.0177 \times \text{SL} - 0.0939 \times \text{FS1} + 0.4923
\]

\[
\text{M. ibericus} (n = 42) = +2.38 \pm 0.97 \text{[min, max: 0.49, 4.89]},
\]

\[
\text{[5–95% percentile range: +1.09, +4.01]}
\]

\[
\text{[two syntype workers: +1.95, +3.46, +2.98]},
\]

\[
\text{[5–95% percentile range: −2.19, +0.87, +0.87]}
\]

\[
\text{Queen.} \text{Smallest and only species with microreticulate and not shining surface of anepisternum and katepisternum among all five species.}
\]

3.8.1.2. Redescription. *Worker and queen.* Colour: Head and mesosomal light brown to brownish red, gaster brown to dark brown. Major worker mostly dark brown. Queen dark brown. *Size:* Medium. Queen small, similar in size to major. *Head:* Entirely sculptured with mostly parallel costae, which usually bifurcate above eye level. Sculpture disappears gradually above eye level, at least on vertex. Postocular region on side of head can lack costae, but microsculpture always present. Erect setae abundant on side of head from occiput to mandibular insertion. *Scapus:* Base without lobe, tooth-like processes directed mostly downward. *Labium:*
funicular segment short and flattened, longer than 2nd segment, but shorter than 2nd and 3rd segments together. Clypeus: Median notch well-defined, but shallow in some minor workers. Pronotum: Median costae often arched, middle of pronotum often smooth. Mesonotum: Densely sculptured throughout, costae transverse and mostly parallel. Metapleural: Densely sculptured with fine, irregular costulae with dense punctuation among them, even in minor worker. Propodeum: Mostly rounded; rarely angulated in major worker; smoothly rounded in minor worker. Parallel costae on side. Surface of 1st gaster tergite: Entire surface imbricate.


Male. Not available.

mtDNA (Figs. 2, 3). Six haplotypes of COI (1375 bp) known; GenBank accession numbers: see Table S1.

Life history. Multicoloniality and swarming flights reported from Mainz, Germany (Heller, 1971).

Distribution (Fig. 2; Table S1). Bulgaria, Croatia, France, Germany, Greece, Italy, Romania, Slovenia, Spain, Switzerland.

3.8.2. Messor ponticus sp.n.

(Figs. 2 and Table S1, Supplementary material).

Corresponds to Lineage 2.

Type material as designated hereby: Holotype. BULGARIA BG016 [/] viz. Strouma valley [/] SW Zemen [-] 06.10.2004 [/] leg. T. Ljubomirov [-] "14759" [-] Holotype [/] "Messor" [/] "ponticus" [on the reverse side: Top specimen design. Csósz 2016] (NHMW).

Paratypes: 11 workers labelled as holotype: BULGARIA BG016 [/] viz. Strouma valley [/] SW Zemen [-] 06.10.2004 [/] leg. T. Ljubomirov (5 paratype workers: NHMW; 6 paratype workers: HNHM). Other material examined: See Table S1.

3.8.2.1. Diagnosis. Worker and queen. Usually darker than M. ibericus, M. mcarthuri sp.n. and M. structor. Generally, body more finely sculptured than in all other species, especially head. Discrimination from M. structor and M. muticus by shorter 1st funicular segment, by base of scape without lobe. Discrimination from all other species by reduced number of standing setae on side of head and, in some individuals, by shallower clypeal notch. Mesopleuron more regularly rugose than in all other species. Microsculpture of 1st gastral tergite similar to M. structor and M. muticus but clearly less imbricate than in M. ibericus and M. mcarthuri.

Worker. Generally larger than M. ibericus, M. mcarthuri, and M. muticus, and similarly sized as M. mcarthuri sp.n. Discrimination from all other species by postocular region of head mostly lacking costae and costae, almost smooth and shining in major worker. For individuals difficult to discriminate from M. structor or M. muticus using qualitative morphology, morphometrics-based linear discriminant functions D10ponticus vs. structor and D9muticus vs. ponticus, respectively, are available.

Linear discriminant function D10ponticus vs. structor uses 10 morphometric characters with a classification success of 97.9%:

\[
D_{10ponticus vs. structor} = -0.0068 \times CW - 0.0739 \times FS1 - 0.0464 \times FS3 - 0.0203 \times PPH + 0.0317 \times FS2 + 0.0152 \times NOH - 0.0089 \times NOL - 0.0131 \times PEW + 0.0104 \times MW - 0.0051 \times SL + 6.9091
\]

\[
M. ponticus (n = 46) = \text{mean: +3.42 ± 1.39 [min, max: +0.83, +6.58], [5-95% percentile range: +1.43, +5.33], [type series (n = 6) mean: 2.38 [min, max: +1.64, +3.12]}
\]

\[
M. structor (n = 174) = \text{mean: -0.90 ± 0.87 [min, max: -3.19, +1.45], [5-95% percentile range: -2.67, +0.40], [type: -1.37]}
\]

Linear discriminant function D9muticus vs. ponticus uses nine morphometric characters with a classification success of 97.5%:

\[
D_{9muticus vs. ponticus} = -0.0067 \times CW - 0.0441 \times TLD - 0.0073 \times ML + 0.0495 \times FR + 0.0193 \times Ocl + 0.0232 \times PPH + 0.0103 \times NOL + 0.0591 \times FS1 - 0.0347 \times FS3 - 2.9232
\]

\[
M. ponticus (n = 46) = -2.18 ± 0.84 [min, max: -4.22, -0.68], [5-95% percentile range: -3.56, -0.89], [type series (n = 6) mean: -1.43 [min, max: -2.08, -0.89]}
\]

\[
M. muticus (n = 72) = +1.39 ± 1.09 [min, max: -1.50, +5.07], [5-95% percentile range: -0.46, 3.01]}
\]

Queen. Larger than M. ibericus, similar in size to all other species.

3.8.2.2. Etymology. Named after its distribution area around the Black Sea (Lat.: Pontus Euxinus).

3.8.2.3. Description. Worker and queen. Colour: Dark brown to black, gaster always blackish. Size: Large. Head: Usually finely sculptured with longitudinal, regular costae below eye level, but can be reduced in minor worker; postocular region can lack sculpture entirely, sometimes smooth and shining. Very sparse, short, erect setae on side of head and genae, sometimes lacking almost entirely. Scape: Base without lobe. Laterally directed, tooth-like processus in major worker, less distinct in minor worker. 1st funicular segment short and flattened, longer than 2nd segment, but shorter than 2nd and 3rd segment together. Clypeus: Median notch very shallow, often lacking entirely. Pronotum: Middle smooth and shining; laterally mostly regular, bended, fine costae, can be reduced to shallow microreticulation in minor worker. Mesonotum: Densely sculptured throughout; costae transverse, with punctuation amongst. Mesopleuron: Regular transverse costulae with slightly irregular, well-developed punctuation among them; costulae can be very fine even in major worker. Propodeum: Sometimes angulated even in minor worker, in major worker almost tooth-like, with bended costae. Surface of 1st gaster tergite: Base imbricate, middle smooth and shining with isolated snow-flakes-like structure. Covered with sparse, thick and long, whitish hairs, some of which decumbent or subdecumbent.


Male. Aneupisternum: Microreticulate, but shining. Katepisternum: With longitudinal costulae; shining even if microsculpture among some costulae.

mtDNA (Figs. 2, 3). Three haplotypes of COI (1375 bp) known; GenBank accession numbers: see Table S1.

Distribution (Fig. 2; Table S1). Bulgaria, Romania, Turkey, Ukraine.

3.8.3. Messor structor (Latreille, 1798)

(Figs. 2 and Table S1, Supplementary material).

Corresponds to Lineages 3 + 4 + 5.

Formica structor Latreille, 1798: 46 (worker, male).

List of verified junior synonyms

Messor rudivis Fabricius, 1804) – 2 q, labelled: rudivis – ZMUC – originally described as: Formica rudivis Fabricius, 1804

Messor lapidum Fabricius, 1804) – 1 w (without head), labelled: lapidum – ZMUC – originally described as: Formica lapidum Fabricius, 1804

Type material as designated hereby: Neotype. FRANCE FR020 [/] 300 m E Gignac [/] S of Brive-la Gaillarde [-] (1’28°E/45°W) [/] 01.10.2006 [/] leg. Galkowski [-] Neotype [/] “Messor” [/] “structor” [on the reverse side: Top specimen design. Csósz 2016] (NHMW)

9 additional workers from the same nest series and labelled as the neotype are deposited in NHMW (5 workers) and in HNHM (4 workers). If neotype is destroyed or lost, a replacement neotype can be designated from these 9 workers.
3.8.3.1. Diagnosis. Worker and queen. Mostly lighter than *M. ponticus* sp.n. and *M. muticus* sp.n. Discrimination from *M. ibericus*, *M. ponticus*, and *M. mcarthuri* sp.n. by long 1st funicular segment, by base of scape with rounded lobe, and by coarser sculpture of mesopleuron, from *M. ponticus* by abundant standing setae on side of head. Head more irregularly costate than in *M. mcarthuri*. Surface of 1st gastric tergite less imbricate than in *M. ibericus* and *M. mcarthuri*, similarly to *M. ponticus* and *M. muticus*.

Worker. Similarly sized as *M. ibericus* and *M. muticus*, smaller than *M. ponticus* sp.n. and *M. mcarthuri* sp.n. For individuals difficult to discriminate from *M. ibericus*, *M. ponticus*, or *M. muticus* using qualitative morphology, linear discriminant functions based on morphometrics available; discrimination from *M. ibericus*: D2 *ibericus* vs. *structor* and D10 *ponticus* vs. *structor* in Diagnosis of *M. ibericus* and *M. ponticus*, respectively; discrimination from *muticus*: D12 *muticus* vs. *structor* below.

Linear discriminant function D12 *muticus* vs. *structor* uses 12 morphometric characters with a classification success of just 87.0%, but as far as currently known, *M. structor* and *M. muticus* are well separated longitudinally (Table S1) and ecologically (Fig. 4); future classification success may improve when using a revised character set:

\[
\begin{align*}
D_{12_{\text{muticus vs. structor}}} &= +0.0428 \times FS1 - 0.0617 \times FS3 + 0.0430 \times OC \times 0.0159 \times PPW + 0.0253 \times NOH - 0.0518 \times TLD + 0.0336 \times FR + 0.0048 \times CW - 0.0062 \times SL + 0.0230 \times FS2 - 0.0163 \times PPH + 0.0066 \times NOL - 1.2192 \\
M. \text{structor} (n = 174) &= \text{mean: } -0.64 \pm 0.95 \text{ [min, max: } -3.05, +2.97\text{], [5–95% percentile range: } -2.12, +0.99\text{, [neotype: } -0.52\text{]} \\
M. \text{muticus} (n = 72) &= \text{mean: } 1.55 \pm 1.11 \text{ [min, max: } -1.44, +4.77\text{], [5–95% percentile range: } -0.24, +3.42]\end{align*}
\]

Queen. Larger than *M. ibericus*, similarly sized as the other three species. Discrimination from *M. ibericus* by reduced reticulation and rugosity of anepisternum and katepisternum.

3.8.3.2. Redescription. Worker and queen. Colour: Head and mesosoma brown to brownish red, gaster brown; major workers sometimes entirely dark brown. Size: Medium. Head: Mostly costate entirely. Parallel costae below eye level. Especially in major and medium worker, vertex or even entire surface above eye level with irregular costae and costulae with microreticulation amongst. In minor and sometimes in media worker, costae reduced in postocular region; head almost smooth and frequently shining, with only fine costae and microreticulate. Erect setae abundant on side of head from occiput to mandibular insertion. Scape: Base with rounded, laterally-downward directed lobe, very distinct in minor worker; can be reduced in major worker. 1st funicular segment long and flattened, almost as long as 2nd and 3rd segment together. Clypeus: Median notch well defined; can be shallow in minor worker. Pronotum: Middle mostly coarsely and irregularly costate throughout. Mesonotum: Coarsely and irregularly costate throughout. Mesopleuron: Coarsely sculptured with irregular transverse costae, with shallow punctuation amongst. Minor workers less costate with almost no punctuation. Propodeum: Mostly smoothly rounded; often obliquely rounded or rarely angled in major workers. Surface of 1st gaster tergite: Basis imbricate, middle with isolated snow-flakes-like structure. Covered with yellowish, thin, sparse hairs, some of which decumbent and subdecumbent in some individuals.


\[\text{mitDNA (Figs. 2, 3). Thirty-one haplotypes of COI (1375 bp) known; GenBank accession numbers: see Table S1.}\]


(Figs. 2 and 5ad, bd, cd, dd, ed, fd; Table S1, Supplementary material).


Other material examined: See Table S1.

3.8.4.1. Diagnosis. Worker and queen. Colour similar to *M. ponticus* sp.n., usually darker than *M. ibericus*, *M. mcarthuri* sp.n., and *M. structor*. Discrimination from *M. ibericus*, *M. mcarthuri*, and *M. ponticus* by base of scape with rounded lobe, long 1st funicular segment, and by coarser sculpture of mesopleuron, from *M. ponticus* by abundant standing setae on side of head. Surface of 1st gastric tergite less imbricate, than in *M. ibericus* and *M. mcarthuri*, similarly to *M. structor* and *M. ponticus*.

Worker. Similarly sized as *M. ibericus* and *M. structor*, smaller than *M. mcarthuri* sp.n. and *M. muticus* sp.n. For individuals difficult to discriminate from *M. ponticus* or *M. structor* using qualitative morphology, morphometrics-based linear discriminant functions D9 *muticus* vs. *ponticus* and D12 *muticus* vs. *structor* available in Diagnosis of *M. muticus* and *M. structor*, respectively; classification success using D12 *muticus* vs. *structor* is just 87.0%, but as far as currently known, *M. structor* and *M. muticus* are well separated longitudinally (Table S1) and ecologically (Fig. 4).

Queen. Larger than *M. ibericus*, similarly sized as *M. mcarthuri* sp.n., *M. ponticus* sp.n., and *M. structor*. Discrimination from *M. ibericus* by reduced reticulation and rugosity of anepisternum and katepisternum.

3.8.4.2. Redescription. Worker and queen. Colour: Head and mesosoma brown to brownish red, gaster dark brown. Major workers can be dark brown entirely. Size: Medium sized. Head: Mostly entirely costate. In minor and sometimes in media worker, costae reduced in postocular region; head almost smooth and frequently shining, with fine costae and microreticulate, frequently shining. Erect setae abundant on side of head from occiput to mandibular insertion. Scape: Base with rounded, laterally-downward directed lobe; very distinct in minor worker; can be reduced in major worker. 1st funicular segment long and flattened, almost as long as 2nd and 3rd segment together. Clypeus: Median notch well defined; can be shallow in minor worker. Pronotum: Middle mostly coarsely and irregularly costate throughout. Mesonotum: Coarsely and irregularly costate throughout. Mesopleuron: Coarsely sculptured with irregular transverse costae; shallow punctuation amongst costae. Minor workers less costate with almost no punctuation. Propodeum: Mostly smoothly rounded; in major worker, often obliquely rounded, rarely angled. Surface of 1st gaster tergite: Basis imbricate, middle with isolated snow-flakes-like structure. Covered with yellowish, thin, sparse hairs, some of which decumbent and subdecumbent in some individuals.


\[\text{mitDNA (Figs. 2, 3). Ten haplotypes of COI (1375 bp) known; GenBank accession numbers: see Table S1.}\]
**Distribution** (Fig. 2; Table S1). Armenia, Kazakhstan, Kyrgyz Republic, Romania, Russian Federation, Ukraine.

3.8.5. **Messor mcarthuri** sp.n.  
(Figs. 2 and Sae, be, ce, de, ee; Table S1, Supplementary material). Corresponds to Lineage 7.  
*Other material examined*: See Table S1.

3.8.5.1. Etymology. Named in honour of ant taxonomist Archie J. McArthur who we remember for his outstanding enthusiasm, dedication, and warmth.

3.8.5.2. Diagnosis. **Worker and queen**. Colour generally lighter as in *M. muticus* and *M. ponticus* sp.n. Discrimination from all species based on very regular costate sculpture of head and body generally, petiole included. Discrimination from *M. muticus* and *M. structor* by base of scape lacking lobe, short 1st funicular segment, from *M. muticus*, *M. ponticus*, and *M. structor* by imbricate surface of 1st gaster tergite, and from *M. ponticus* by side of head covered abundantly with standing setae.

**Worker.** Similarly sized as *M. ponticus*, usually larger than *M. ibericus*, *M. muticus*, and *M. structor*. Safe discrimination from *M. structor* only feasible by morphometrics. Discrimination from all other species based on the regular costate surface of body including petiole and postpetiole. Also using morphometrics, *M. mcarthuri* sp.n. can be separated safely from any other species treated in this revisionary using the 1st root of the linear discriminant function D15 in Table 1.

**Queen.** Larger than *M. ibericus*, similarly sized as *M. muticus*, *M. ponticus* sp.n., and *M. structor*. Discrimination from *M. ibericus* by reduced reticulation and rugosity of anepisternum and katepisternum.

3.8.5.3. Description. **Worker and queen**. Colour: Head and mesosoma brown to brownish red, gaster dark brown. Major workers can be dark brown entirely. **Size**: Large. **Head**: Regularly costate throughout, interstices reticulate. In major and most medium workers, costae present on much of postocular region. Erect setae abundant on side of head from occiput to mandibular insertion. **Scape**: Base without lobe. Laterally directed, tooth-like process in major worker present, less distinct in minor worker. 1st funicular segment short and flattened, longer than 2nd segment, but clearly shorter than 2nd and 3rd segment together. **Clypeus**: Median notch well defined; can be shallow in minor worker. **Pronotum**: Mostly shining, laterally irregularly running costae. **Mesonotum**: Regularly costate entirely. **Anepisternum**: With slightly irregular transverse costae, well developed punctuation among costae. **Surface of 1st gaster tergite**: Entire surface imbricate. Covered with thick whitish, sparse hairs, some of which decumbent or subdecumbent. **Katepisternum**: Smooth and shining. **Metanotum**: Regularly costate. **Aneupisternum**: Smooth and shining. **Mesoraceme**: Terminal with fine costulae, middle smooth and shining.

**mtDNA** (Figs. 2, 3). Six haplotypes of COI (1375 bp) known; GenBank accession numbers: see Table S1.  
*Distribution* (Fig. 2; Table S1). Greece, Turkey.

4. Conclusion  
Here, we used a cohesive protocol for integrative taxonomy to explore species limits in the harvester ant *Messor* "structure", a nominal species with an outstanding number of subspecies and junior synonyms. We used information from TM, Wolbachia symbionts, mtDNA, AFLP, and ENM and discovered that *Messor* "structure" consists of five species, that is, *M. structor*, *M. muticus*, and *M. ibericus* and the two new species *Messor ponticus* sp.n. and *M. mcarthuri* sp.n. Importantly, resolving incongruences among data sets unravelled possible, complex evolutionary histories now open for further research evolution. We note that in the case of *M. "structure"* the final species limits would not have been retrievable via any single data set. The newly delimited species differ in their distribution area and ecology, and likely beyond, highlighting how important it is to understand evolutionary processes for drawing correct conclusions on relevant species’ attributes.

5. Data accessibility  
GenBank accessions for mitochondrial DNA sequences, detailed information of sampled specimens on sampling location and further details on the integrative-taxonomic procedure are available as Supplementary material. Alignments of the mitochondrial sequences, the TM and AFLP raw data, as well as the climate, soil and vegetation layers and the layer correlation matrix used in ENM are available as doi: https://doi.org/10.5061/dryad.mj43d20.

6. Author contributions  
F.M.S., S.C., B.M., C.S., W.A., and B.C.S.-S. designed the study. F.M.S. and B.C.S.-S. coordinated the study. F.M.S., S.C., B.M., and B.C.S.-S. did the field work. C.F., S.H., W.A., and B.C.S.-S. performed the molecular experiments. S.C. performed the morphometric analyses, with contributions by F.M.S. B.M. performed the qualitative-morphological analyses. A.G. did the ecological niche modelling, with contributions by F.M.S. and L.R. F.M.S., S.C., A.G., S.H., W.A., and B.C.S.-S. analysed the data. F.M.S. wrote the draft manuscript with contributions from S.C., B.M., A.G., W.A., and B.C.S.-S., which was revised by all authors. All authors read and approved the final version of the manuscript.

Acknowledgements  
For collecting *Messor* samples, to all persons listed in Table S1; for help in the molecular laboratory, to Philipp Andesner, Elisabeth K. Zangerl, and Raphael Strohmaier; for help with GIS analyses, to Georg Leitinger and Raphael Strohmaier; for advice regarding Russian literature on *Messor* and, generally, morphological features, to Alexander G. Radchenko; for constructive criticism of an earlier manuscript version, to two anonymous reviewers. A.G. was funded by the Austrian Science Fund (P 23409), B.M. by several Synthesys EU FP6 grants and the Bolyai János Scholarship of the Hungarian Academy of Sciences.

Conflicts of interest  
None.

Appendix A. Supplementary material  
Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2018.04.005.


