Vector-GMA: A New Easy Method for Analyzing Morphometric Data

Dmitry S. Kopylov (✉ aeschna@yandex.ru)

A.A. Borissiak Palaeontological Institute, Russian Academy of Sciences

Research Article

Keywords: geometric, morphometry, GMA, landmark, phylogeny, morphology, Ichneumonidae, Cretaceous, fossil, insect.

Posted Date: July 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-702822/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Vector-GMA: a new easy method for analyzing morphometric data

Dmitry S. Kopylov

A.A. Borissiak Palaeontological Institute, Russian Academy of Sciences, Moscow 117647, Russia

Cherepovets State University, Cherepovets, 162600, Russia

aeschna@yandex.ru *ORCID: <https://orcid.org/0000-0003-2013-544X>*

Abstract

The new method of geometric morphometric analysis, called Vector-GMA is proposed. Unlike the classic TPS-based GMA, the new method allows considering missing and inapplicable landmarks and analyzing articulated structures. It also contains a set of tools for analyzing particular morphological features and searching for promising identification characters. The results of the analysis can be easily converted into phylogenetic matrix and analyzed with phylogenetic software. For the case study, the fore wings of all Cretaceous ichneumonids are analyzed.

Keywords: geometric, morphometry, GMA, landmark, phylogeny, morphology, Ichneumonidae, Cretaceous, fossil, insect.

Introduction

Morphometry is one of the most important components of any study of biological object morphology. The comparison of metric parameters (segment lengths, angles, etc.) allows us to obtain important information that is used in taxonomic, morphology, and ecology research (Štarhová Serbina & Mennecart 2018; Wilson *et al.* 2015). Often, researchers use a small number of morphometric parameters, identified mainly in an intuitive way. But how fully do we use this data? For example, take the Ichneumonidae wing, a fairly simple object for morphometric analysis. On this wing, we can easily distinguish about $n = 20$ homologous landmarks, built at the points of the veins branching. It is easy to calculate that $N_1 = C_2^n = 190$ simple measurements, and $N_2 = C_2^{N_1} = 17955$ pairwise comparisons of the lengths of the segments can be made between these landmarks. Given that in ordinary taxonomic works, rarely more than a few dozen morphometric parameters are used, it is easy to see how incomplete we use morphometric information.

One of the effective solutions to this problem is GMA: Geometric Morphometric Analysis. In this method, the objects defined by the landmark map are represented as a set of thin plates. Further, by deforming these plates, the strain energies and Procrustean distances are calculated: numerical parameters reflecting the degree of differences between each pair of samples. Using the UPGMA

Procrustean distance matrix, we can build a relationships dendrogram for the sample, which can be used with a certain degree of caution as a proxy for reconstructing phylogenetic relationships within the sample. The GMA method also allows performing a number of valuable analyses that allow identifying or confirming the clustering of objects in the sample (Principal Components, Canonical Variables). An important component of the method is transformational grids – a method for visualizing the deformation of thin plates (Rohlf 2015).

The TPS GMA method is a very powerful tool for morphometric analysis, but a number of disadvantages limit the scope of its application:

1. The necessity of mapping all landmarks for each sample in the set is perhaps the most significant Achilles' heel of the method. First, not all samples may have a complete set of homologous points (for example, some species may have some of the vein furcations lost). Secondly, in some cases, we have to deal with not fully preserved samples (which is crucial for the application of the method in paleontology).
2. In GMA method, the object is analyzed as a whole. Thus, we receive only a general conclusion about the objects clustering, but we can not specify the parameters contributed to the result. While in morphological studies, the main task is usually in finding specific distinguishing features.
3. GMA does not allow analyzing articulated structures (arthropod legs, skeletons, fern fronds, etc.).
4. The results of GMA can be hardly implemented to phylogenetic analysis (Catalano *et al.* 2010; Palci & Lee 2019).
5. GMA mathematical algorithms are extremely complex.

To solve the problems associated with the use of GMA, an alternative approach is proposed here. Instead of analyzing the object as a whole, I suggest using the landmarks to build a network of vectors and then analyze them. This approach allows us to flexibly work with missing landmarks, identify the most significant morphometric parameters, and analyze articulated structures. The results of this analysis are easily incorporated into phylogenetic matrixes, and the mathematical model used here is reduced to the simplest vector algebra. The proposed method is called Vector-GMA (VGMA).

As the case study, the Cretaceous ichneumonids are analyzed here with the proposed method.

Material and methods

The proposed method is implemented in the form of program VGMA (Windows only, Supplement 1). The program is written in C#, Windows Forms, .Net Framework 4.0, Microsoft Visual Studio 2019 (open source code for version 0.2.2 alpha is in Supplement 2). The Quick Guide to VGMA is available as Supplement 3.

In the case study, all Cretaceous species of Ichneumonidae are analyzed, excluding four poorly preserved ones. In addition to the valid species, I include here five specimens from Burmese amber, which represent five new species of Novichneumoninae (Kopylov *et al.*, in press). For the analysis, I used original photos, sketch drawings and figures from taxonomic papers (see Table 1 for references). To check the reproducibility of the results, *Khasurtella buriatica* was included into the calculation twice:

one specimen is the holotype and another is paratype. The complete list of Cretaceous ichneumonid species is given in the Table 1. In this study, only fore wings were analyzed. The source data for the analysis, VGMA map and phylogenetic analysis files are given in the Supplement 4.

Phylogenetic analysis was performed in PAUP* 4.0a (build 169). For the majority of the analysis Heuristic search with the unlimited number of possible max-parsimonious trees has been used. Only in Single Angles analysis, the max-trees setting has been reduced to 1000 to overrun memory overflows. For reweighting the characters the integrated function of PAUP has been used. The reweighting has been conducted according to Rescaled consistency index (RC). All PAUP analyses logs and retained trees are given in the Supplement 5.

Results

Morphometric data digitizing

The method is intended for the analysis of flat or almost flat objects (insect wings, plant leaves, fish scales, etc.). The object digitizing is performed on the basis of images (photos, exact drawings).

First, landmark mapping is performed, according to the same principles as in GMA. Just like in GMA, all landmarks must be plotted on the object. However, in the VGMA method, it is possible to set the Missing (lost due to preservation) or Inapplicable (the landmark has no sense in a particular case) parameter for the selected landmarks. The position of these landmarks on the object does not matter.

In the next step, a network of vectors is superimposed on the landmark field. Basically, it is convenient to plot vectors in accordance with anatomical structures (wing or leaf veins, scale contour, etc.). However, the user is free to build vectors on any pair of landmarks at his discretion, depending on the particular study task. Vectors relying on Missing or Inapplicable landmarks automatically get the corresponding property. The user can also set the Inapplicable property for the vector manually.

For more information on using Digitizer, see the Quick Guide (Supplement 3). When the objects are digitized, the user starts an analysis that performs sample alignment and aligned vector comparison.

Vector comparison

This simple instrument allows visualization of comparative scalar lengths of the selected vectors among the sample. The comparative lengths of vectors are shown in color scale red-black-green for short-medium-long sectors Fig. 1.

Pairwise Analysis

This instrument provides a set of tools for pairwise vector comparison using scalar lengths, angles and vector differences. In “Scalars” mode the main analyzing parameter is the ratio between scalar lengths of the vectors, in “Angles” – the angle between the vectors, in “Vectors” – the vector differences between the vectors (scalar length and angle of the resulting vector). The program generates the data grid representing all pairs of vectors, and gives the statistics on the pairwise comparison, including number of applicable specimens, minimal and maximal values of analyzed parameters, the ranges,

maximal/minimal ratios, standard deviations etc. Sorting the data table with these parameters allows easy search for the most significant morphometric characters. The program also allows generating the plots graphically representing the character distribution among the sample.

For example, scalar length comparison of the veins 1-Cu and 2+3-M in Cretaceous Ichneumonidae reveals good clustering: subfamilies Tanychorinae and Novichneumoninae are clearly separated from other representatives, Novichneumoninae and Labenopimplinae overlapping, however, the mathematical expectations of the samples are clearly different (Fig. 2). This example demonstrates how the VGMA method allows us finding a significant morphometric taxonomic character that may be used to identify subfamilies. It is worth noting that the mentioned veins are located at a considerable distance, and with the classical "intuitive" approach to taxonomic morphometry, this character would hardly be revealed.

The method of angular comparison is generally less informative than the scalar one. However, it also allows finding significant characters. For example, the angle between 2-Rs and 2+3-M may be used as a secondary character for subfamily identification (Fig. 3): the angle is strongly acute in Tanychorinae, usually 45° to 90° in Palaeoichneumoninae and Labenopimplinae and over 90° in Novichneumoninae. However, we can see that the values for different subfamilies are overlapping, so this character may be used as a hint, but not an unambiguous feature.

The vector subtraction results (Fig. 4) may be the most confusing for interpretation. First, the result of vectors subtraction is a vector, so it can't be completely represented with a single numeric value. That makes further operations with the results rather complicated. Second, vector subtraction is the comparison method hardly suitable for identification keys and taxa diagnoses. So, it may be cautiously used for sample clustering, but has no great practical sense.

Converting the morphometric data to phylogenetic matrix

The program implements the conversion of morphometric data into the popular in phylogenetics Nexus format (testing was carried out in conjunction with PAUP* 4.0). This conversion is essentially a special case of converting the ratio scale values to rank scale. The conversion is performed by dividing the ratio scale into n segments and then assigning each segment a certain rank. In this case, I use $n = 10$ ranks (0, 1 ... 9), as such values are most convenient for phylogenetic coding. The conversion methods used in VGMA are shown in Fig. 5. The conversion methods used in VGMA are not original and have already been described earlier (Garcia-Cruz & Sosa 2006; Wiens 2001).

In the "0-MAX" method, maximal rank 9 is assigned to the maximal value of the sample, rank 0 – to 0. When no samples fit into the lower ranks, they may be removed. The method "-PI - +PI" is the analogue of "0-MAX" for angular metrics.

In the "MIN-MAX" method, rank 0 is assigned to the minimal value of the sample. This method is more sensitive than "0-MAX". However, in the cases with all values in the sample are close, this sensitivity may strongly overweight the insignificant differences. Thus, methods "0-MAX" and "-PI - +PI" are generally more appropriate than "MIN-MAX".

"Log MIN-MAX" method is recommended for comparative metrics, such as pairwise scalar ratios. It differs from "MIN-MAX" in using logarithmic scale instead linear one.

Composing the phylogenetic dataset

VGMA allows composing phylogenetic matrixes using six types of data. For all analyses, only aligned shapes are used, which means absolute dimensions of analyzed objects do not matter, but only relative parameters of the vectors.

The implemented data are:

1. Vectors Applicability. Coded "0" for applicable vectors, "1" for inapplicable, and "?" for missing ones. Unlike other characters used in the VGMA model, Vectors Applicability is a true qualitative character, very similar to classic characters used in morphological phylogenetic analysis. In the logic of VGMA, an inapplicable vector corresponds with an undeveloped morphological structure. In other words, applicable vector meaning existing character and vice versa.
2. Single Scalars. The analyzed values are scalar lengths of homologous vectors among the sample. The most convenient conversion method is "0-MAX".
3. Single Angles. The analyzed values are vector angles of homologous vectors among the sample. The reference angle is the angular average of all analyzed vectors. The most convenient conversion method is "-PI - +PI".
4. Scalar Ratios. The analyzed values are pairwise ratios of scalar vector lengths. The most convenient conversion method is "Log MIN-MAX".
5. Pairwise Angles. The analyzed values are pairwise angular differences between vectors. The most convenient conversion method is "-PI - +PI".
6. Vector Differences. The analyzed values are scalars of vector subtractions. The most convenient conversion method is "0-MAX".

For the first three data types, the number of generated characters equals n = the number of vectors. For the latter three - $N = C_2^n$.

VGMA allows using a single data type for composing the matrix, as well as combining the mentioned data types in any combinations. The phylogenetic matrix produced by VGMA may be used in phylogenetic analysis software as it is. However, the users are free in expanding the dataset manually with their own data.

The problem of character weights

One of the problems in converting morphometric data into the phylogenetic matrix is the character weights. This issue does not cause big problems while working with a single data type, but combining different data types in one matrix requires character weight consideration.

First, pairwise data types return much more characters than single ones. That means analyzing these characters with the same weight could badly underweight the single ones.

Second, quantitative characters in this model are coded as ordered with nine possible steps (0 to 9), while Vector Applicability data type is unordered with only "0" and "1" values (one step). This makes quantitative characters nine times "stronger" than the qualitative one.

A satisfactory solution to this problem is beyond the scope of this work. I will be glad to hear colleagues' suggestions on this issue. At this stage, the problem is solved based on the assumption that all data types in the matrix should have a similar weight. The weight "1" is assigned for the Vector Applicability characters as the only qualitative ones. Then, all quantitative character weights are divided by 9 to balance the number of steps. Then, the pairwise character weights are also divided by $\frac{C_2^n}{n}$ to balance the differences in the character numbers.

The assumption underlying this method is quite speculative, but it allows balancing the characters. It also does not take into account potential differences in the character weights within the data type. The program allows the users to set the character weights manually.

An additional method to improve the results of phylogenetic analysis of VGMA matrices is to use the RC Reweight characters function, implemented, for example, in PAUP* 4.0.

When supplementing the VGMA matrix with own characters, it is also necessary to consider the problem of the character weights.

Case study: Ichneumonidae of Mesozoic

Ichneumonids were chosen as an object for the case study due to the ambiguity of the applicability of morphometry in their taxonomy. For example, representatives of the subfamily Tanychorinae have a very peculiar venation, which allows us to hope for good applicability of morphometry for their taxonomic isolation from other representatives of the family. On the contrary, the subfamilies Palaeoichneumoninae and Labenopimplinae have very similar venation of the fore wing and differ mainly in body structure characters. The separation of these subfamilies should become a serious challenge for the proposed method. For example, in the previous work (Li *et al.* 2019), we were unable to separate the representatives of these subfamilies using the GMA - UPGMA method. Finally, the Burmese ichneumonids have very different types of venation. Novichneumon and Caloichneumon have a unique type of venation within Cretaceous representatives of the family. The wing of Heteropimpla rather resembles the wings of Labenopimplinae. The five not yet described representatives of the Burmese fauna included in the analysis are also very diverse in venation.

The samples were digitized with 22 landmarks and 37 vectors (Tables 2, 3, Supplement 4).

The results were evaluated in accordance with traditional ideas about the phylogeny of Mesozoic ichneumonids (Kopylov 2009, 2010; Li *et al.* 2017, 2019; Rasnitsyn 1980).

Tanychorinae is the paraphyletic subfamily, traditionally assigned to Ichneumonidae. Tanychorines are ancestral (or sister-group) to all other ichneumonids and family Braconidae. The most plesiomorphic within Tanychorinae is the genus *Amplicella*, while the genera *Paratanychora* and *Tanychorella* are the most apomorphic ones.

Palaeoichneumoninae is the paraphyletic subfamily, most probably descending from Tanychorinae and probably ancestral to Labenopimplinae. The genus *Rudimentifera* has some plesiomorphies in the fore wing venation which make it rather close to the higher tanychorines.

Labenopimplinae is a highly diverse and probably polyphyletic group. It includes genera with very archaic venation (*Ramulimonstrum*, *Armanopimpla*), with venation similar to Palaeoichneumoninae

(*Labenopimpla*, *Micropimpla*), with pimpline-like venation (*Rugopimpla*), and the array of various small simplified forms, mainly from ambers (*Agapia*, *Agapteron*, *Urotryphon*, *Eubaeus*, *Pareubaeus*, *Catachora*). The genus *Heteropimpla*, described in Labenopimplinae, is more likely to be placed to Novichneumoninae (Kopylov *et al.* in press).

Novichneumoninae is the endemic Burmese subfamily. It has been originally described by two species with unique venation (*Novichneumon*, *Caloichneumon*). Now another five species are being described from Burmese amber, all identified as Novichneumoninae, which significantly expands the original diagnosis of the subfamily (Kopylov *et al.* in press).

Mesozoic ichneumonid fauna also includes one species of Labeninae (*Albertocryptus dosseus*) and one not assigned to any subfamily (*Tryphopimpla xoridoptera*).

First, all data types were analyzed separately.

Vector Applicability data retains the highly-polytomic cladogram (Fig. 6). Reweighting the characters does not affect the results seriously (Fig. 7). In our data set, this method allows only the quite rude separation of Tanychorinae from other ichneumonids, while the relations between other subfamilies are not revealed.

Single Scalars data allows distinguishing Tanychorinae from other subfamilies (Fig. 8). The majority of Palaeoichneumoninae species are placed between Tanychorinae and Labenopimplinae, which corresponds with their traditional position (Kopylov 2009, 2010). The majority of Burmese species are grouped together, but with minor errors: *E. abdominalis* are united with Burmese species, while the specimen BA01 18114 is not included. The separation of genera within subfamilies are quite poor. Reweighting the characters doesn't improve the result (Fig. 9).

Single Angles data returns unsatisfactory results, neither with and without character reweighting (Figs. 10, 11). In both analyses this data type couldn't even complete the most simple task of isolating Tanychorinae from other ichneumonids.

Scalar Ratios data retains the most sufficient result among the single data type analyses (Fig. 12).

Burmese species come into one group, excluding two, placed within Tanychorinae.

Palaeoichneumoninae are placed between Tanychorinae and Labenopimplinae, excluding two errors: *P. mirabilis* is placed within Micropimpla and *R. intermedium* is placed within Palaeoichneumoninae. The second error is quite expectable: *R. intermedium* has very peculiar venation, rather similar with palaeoichneumonines. Reweighting the characters doesn't improve the results (Fig. 13).

Pairwise Angles data retains the most poor results. This method even can't distinguish Tanychorinae from other subfamilies (Figs. 14, 15).

Vector Differences data displays ambiguous results (Fig. 16). It distinguishes Tanychorinae from others very well. Novichneumoninae are also grouped together, but with *Eubaeus* within the group.

Palaeoichneumoninae and Labenopimplinae are not distinguished. Reweighting the characters doesn't change the results significantly (Fig. 17).

As we can see, no single data type analysis retains completely successful result. Character reweighting also doesn't significantly improves the results. However, combining different data types allows building

better cladograms. For combined analyses the best data types from the single analyses were chosen. One test has been also conducted with all data types included.

All data type analysis returns mediocre results (Fig. 18). Tanychorinae are separated from other subfamilies, but Palaeoichneumoninae, Labenopimplinae and Novichneumoninae are interlacing. Triple character reweighting slightly improves the result: just all subfamilies are grouped correctly with only genera *Ramulimonstrum* and *Labenopimpla* fall into Palaeoichneumoninae (Fig. 19).

The best results were obtained by **combining Vector Applicability + Scalar Ratios, Vector Applicability + Single Scalars + Scalar Ratios** and **Vector Applicability + Scalar Ratios + Vector Differences**. The primary analyses in all cases returned trees with Tanychorinae well separated, but other subfamilies interlacing (Figs. 20, 21, 22). However, reweighting the characters strongly improves the results (Figs. 23, 24, 25). In all three cases re results after reweighting were rather similar.

The best tree has been obtained by using **Vector Applicability + Single Scalars + Scalar Ratios with triple reweighting** (Fig. 24). In this analysis, Tanychorinae are well separated from other ichneumonids. The genera within the subfamily are not well distinguished, however, the most plesiomorphic genera *Amplipella*, *Khasurtella* and *Megachora* are sharing the basal part of the tree, the genera *Tanychora* and *Sinochora* fall into the middle part of the group, and the genera *Tanychorella* and *Paratanychora*, presumably being the most apomorphic tanychorines, take their position as the crown-group of Tanychorinae. The holotype and the paratype of *K. buriatica* are forming one group. The middle section of the tree is occupied with the paraphyletic subfamily Palaeoichneumoninae. *Rudimentifera*, having some tanychorine-like rudiments in the fore wing venation, taking its expected position at the base of palaeoichneumonine stem. The genera *Palaeoichneumon* and *Dischysma* are not completely separated, though it is quite expectable: they have almost similar venation differ only in the number of 2m-cu bullae. However, palaeoichneumonines are split *L. orapa* and *R. intermedium*, belonging to Labenopimplinae. This seems to be the most significant fault of this analysis. For *Ramulimonstrum*, this error seems explicable: this genus has venation quite similar with palaeoichneumonines but differs from them in body structure. For *L. orapa*, this fault is rather strange, taking into account other species of the genus taking their position at the crown of labenopimpline section of the tree. Labenopimplines (excluding two genera mentioned above) forming one group. Micropimpla takes the basal position within the subfamily. Other genera forms three groups: *Rugopimpla* + *Agapia* + *Agapteron*, *Eubaeus* + *Pareubaeus*, *Labenopimpla* + *Catachora* + *Armanopimpla*. The genus *Urotryphon* has been split between the first two groups. *Tryphopimpla*, Labeninae and Novichneumoninae originate from Labenopimplinae. Novichneumoninae (including *Heteropimpla* and the five new species) form monophyletic group.

It should be noted that this method doesn't provide good bootstrap supports for the obtained trees. In the essential part of the topology, only the transition from Tanychorinae to other ichneumonids has good supports. In other cases, the supports beyond 50% are sometimes occurred at the terminal branches of the tree.

Discussion

The Vector-GMA method was developed as an alternative to the classical GMA analysis. The main reason for the development was the inability to use GMA in the analysis of incompletely preserved paleontological objects. The principles of working with the VGMA method are largely borrowed from

GMA and are based on similar landmark mapping. However, the mathematical model underlying VGMA differs significantly from GMA. The proposed method does not pretend to be a full-fledged replacement for GMA, but it allows bypassing some of its limitations. In addition, VGMA offers several tools that can be useful in the work of morphologists and taxonomists.

First of all, VGMA allows using missing and inapplicable landmarks. This makes it possible to include into the analysis not completely preserved objects or objects that have qualitative morphological differences. In the case study, 18 of 65 included specimens have missing landmarks and all of them have inapplicable vectors. However, this did not become an obstacle to the application of the method.

GMA analyzes the object as a whole, which makes it difficult to understand which exactly transformations affect the results. VGMA provides a set of tools for visualization of particular vector transformations (**Vector comparison**, Fig. 1) and for searching the promising identification characters (**Pairwise analysis**, Figs. 2, 3, 4).

VGMA is also suitable for analyzing articulated structures. Not all implemented instruments are applicable for this type of object. However, scalar methods (**Vector comparison**, **Pairwise analysis: Vectors**, **Nexus: Single Scalars and Scalar Ratios**), as well as **Vector Applicability** may be used for analyzing articulated objects.

One of the significant features of VGMA is the possibility to convert morphometric data exactly into phylogenetic matrix. This solves the old problem of implementation GMA data into phylogenetic analysis. In the case study, the Cretaceous ichneumonids were analyzed. For the analysis, only the fore wing venation has been used. I initially understood the insufficiency of wing features for constructing a phylogeny of Cretaceous Darwin wasps. The goal of the experiment was to test how this isolated type of data could help to reconstruct the phylogeny. Thus, the phylogenetic trees presented in this paper should not be considered as true relations within the family, but only as a demonstration of the capabilities of the method. Despite the a priori lack of data, the phylogenetic trees obtained as a result of the analysis quite well coincide with our ideas about the evolution of the family in the Mesozoic. Qualitative features (Vector Applicability) and scalar metrics (Single Scalars and Scalar Ratios) showed the best results. The results obtained here (Fig. 24) are better than the results of a similar analysis using the GMA UPGMA method (Li *et al.* 2019). To build a reliable phylogeny of the family, the matrix obtained using VGMA should be combined with the traditional matrix of morphological data. The idea of combining several morphometric matrices obtained for different parts of the body (for example, for the front and rear wings) also seems potentially promising. Naturally, combining matrices of different types into one will require careful consideration of the weight of the characters. Unfortunately, in this work, the problem of the character weights is solved quite conditionally, while combining different data types makes this problem even harder. At the moment, I cannot give unambiguous recommendations on the reweighting of the characters when combining matrices, and I hope for a discussion from specialists experienced in phylogenetic analysis. The option of combining various data using Bayesian phylogeny models looks potentially interesting.

Acknowledgements

I am grateful to the team of the Arthropod laboratory of the Paleontological Institute, Russian Academy of Sciences for discussing the proposed method. I am grateful to Microsoft Corporation, which provides a wonderful Visual Studio development environment to amateur programmers for free.

Funding

This work has been supported by the Russian Science Foundation, RSF 21-14-00284.

References

- Ashmead, W.H. (1900) Classification of the ichneumon flies, or the superfamily Ichneumonoidea. *Proceedings of the United States National Museum*, 23, 1–220.
- Catalano, S.A., Goloboff, P.A. & Giannini, N.P. (2010) Phylogenetic morphometrics (I): the use of landmark data in a phylogenetic framework. *Cladistics*, 26 (5), 539–549.
<https://doi.org/10.1111/j.1096-0031.2010.00302.x>
- Garcia-Cruz, J. & Sosa, V. (2006) Coding Quantitative Character Data for Phylogenetic Analysis: A Comparison of Five Methods. *Systematic Botany*, 31 (2), 302–309.
<https://doi.org/10.1600/036364406777585739>
- Kopylov, D.S. & Zhang, H. (2015) New ichneumonids (Insecta: Hymenoptera: Ichneumonidae) from the Lower Cretaceous of north China. *Cretaceous Research*, 52, 591–604.
<https://doi.org/10.1016/j.cretres.2014.03.012>
- Kopylov, D.S. (2009) A new subfamily of ichneumonids from the Lower Cretaceous of Transbaikalia and Mongolia (Insecta: Hymenoptera: Ichneumonidae). *Paleontological Journal*, 43 (1), 83–93.
<https://doi.org/10.1134/S0031030109010092>
- Kopylov, D.S. (2010a) Ichneumonids of the subfamily Tanychorinae (Insecta: Hymenoptera: Ichneumonidae) from the Lower Cretaceous of Transbaikalia and Mongolia. *Paleontological Journal*, 44 (2), 180–187. <https://doi.org/10.1134/S0031030110020097>
- Kopylov, D.S. (2010b) A new subfamily of ichneumon wasps (Insecta: Hymenoptera: Ichneumonidae) from the Upper Cretaceous of the Russian Far East. *Paleontological Journal*, 44 (4), 422–433.
<https://doi.org/10.1134/S003103011004009X>
- Kopylov, D.S. (2011) Ichneumon wasps of the Khasurty locality in Transbaikalia (Insecta, Hymenoptera, Ichneumonidae). *Paleontological Journal*, 45 (4), 406–413.
<https://doi.org/10.1134/S0031030111040058>
- Kopylov, D.S. (2012) New Ichneumonidae (Hymenoptera) from the Upper Cretaceous ambers of the Taimyr Peninsula. *Paleontological Journal*, 46 (4), 383–391.
<https://doi.org/10.1134/S0031030112040041>
- Kopylov, D.S. (2018) New Ichneumonids (Hymenoptera: Ichneumonidae) from the Lower Cretaceous Khasurty Locality. *Paleontological Journal*, 52 (3), 284–289.
<https://doi.org/10.1134/S003103011803005X>

- Kopylov, D.S., Brothers, D.J. & Rasnitsyn, A.P. (2010) Two New Labenopimpline Ichneumonids (Hymenoptera: Ichneumonidae) from the Upper Cretaceous of Southern Africa. *African Invertebrates*, 51 (2), 423–430. <https://doi.org/10.5733/afin.051.0211>
- Kopylov, D.S., Zhang, Q. & Zhang, H. (in press) The Darwin wasps (Hymenoptera: Ichneumonidae) of Burmese amber. *Palaeoentomology*.
- Li, L., Kopylov, D.S., Shih, C. & Ren, D. (2017) The first record of Ichneumonidae (Insecta: Hymenoptera) from the Upper Cretaceous of Myanmar. *Cretaceous Research*, 70, 152–162. <https://doi.org/10.1016/j.cretres.2016.11.001>
- Li, L., Shih, P.J.M., Kopylov, D.S., Li, D. & Ren, D. (2019) Geometric morphometric analysis of Ichneumonidae (Hymenoptera: Apocrita) with two new Mesozoic taxa from Myanmar and China. *Journal of Systematic Palaeontology*, 18 (11), 931–943. <https://doi.org/10.1080/14772019.2019.1697903>
- McKellar, R.C., Kopylov, D.S. & Engel, M.S. (2013) Ichneumonidae (Insecta: Hymenoptera) in Canadian Late Cretaceous amber. *Fossil Record*, 16 (2), 217–227. <https://doi.org/10.1002/mmng.201300011>
- Palci, A. & Lee, M.S.Y. (2019) Geometric morphometrics, homology and cladistics: review and recommendations. *Cladistics*, 35 (2), 230–242. <https://doi.org/10.1111/cla.12340>
- Rasnitsyn, A.P. (1975) Vysshie pereponchatokrylye mezozoya [Hymenoptera Apocrita of Mesozoic]. Nauka, Moscow. 134 pp. [In Russian]
- Rasnitsyn, A.P. (1980) *Proiskhozhdenie i evoliutsiia pereponchatokrylykh nasekomykh [The origin and evolution of Hymenoptera]*. Nauka, Moscow. 192 pp. [In Russian]
- Rohlf, F. (2015) The tps series of software. *Hystrix, the Italian Journal of Mammalogy*, 26 (1). <https://doi.org/10.4404/hystrix-26.1-11264>
- Štarhová Serbina, L. & Mennecart, B. (2018) Evolutionary pattern of the forewing shape in the Neotropical genus of jumping plant-lice (Hemiptera: Psylloidea: Russelliana). *Organisms Diversity & Evolution*, 18 (3), 313–325. <https://doi.org/10.1007/s13127-018-0367-5>
- Townes, H. (1973a) Three tryphonine ichneumonids from Cretaceous amber (Hymenoptera). *Proceedings of the Entomological Society of Washington*, 75 (3), 282–287.
- Townes, H. (1973b) Two ichneumonids (Hymenoptera) from the Early Cretaceous. *Proceedings of the Entomological Society of Washington*, 75, 216–219.
- Wiens, J.J. (2001) Character Analysis in Morphological Phylogenetics: Problems and Solutions A. Brower (Ed). *Systematic Biology*, 50 (5), 689–699. <https://doi.org/10.1080/106351501753328811>
- Wilson, L.A.B., Colombo, M., Sánchez-Villagra, M.R. & Salzburger, W. (2015) Evolution of opercle shape in cichlid fishes from Lake Tanganyika - adaptive trait interactions in extant and extinct species flocks. *Scientific Reports*, 5 (1), 16909. <https://doi.org/10.1038/srep16909>
- Zhang, H. & Rasnitsyn, A.P. (2003) Some ichneumonids (Insecta, Hymenoptera, Ichneumonoidea) from the Upper Mesozoic of China and Mongolia. *Cretaceous Research*, 24 (2), 193–202. [https://doi.org/10.1016/S0195-6671\(03\)00031-4](https://doi.org/10.1016/S0195-6671(03)00031-4)

Zhang, J.F. (1991) A new species of *Tanychora* Townes (Hymenoptera, Ichneumonidae). *Acta Palaeontologica Sinica*, 30, 502–504.

Figures

Fig. 1. The example of **Vector Comparison** visualization of some Cretaceous ichneumonid fore wings: A – *Tanychora petiolata*, B – *Palaeoichneumon freja*, C – *Novichneumon longus*. Red lines correspond to veins shorter, than average, and green ones – longer, than average.

Fig. 2. **Pairwise Scalar** visualization of comparative 2+3-M and 1-Cu lengths in Cretaceous ichneumonids.

Fig. 3. **Pairwise Angle** visualization of the angle between 2-Rs and 2+3-M in Cretaceous ichneumonids. Solid lines represent the spread of values within subfamilies, sectors represent the spread excluding outliers.

Fig. 4. **Pairwise Vector** visualization of vector subtraction between pterostigma length and 2-Rs.

Fig. 5. Methods of conversion of morphometric data into phylogenetic code implemented in VGMA.

Fig. 6. Phylogenetic tree retained using only **Vector Applicability** data: strict consensus of 46045 most parsimonious trees. Here and in other trees Tanychorinae are marked with red dots, Palaeoichneumoninae – with green, Labenopimplinae – with blue, Novichneumoninae – with pink, Labeninae – with orange and *Tryphopimpla* – with grey.

Fig. 7. Phylogenetic tree retained using only **Vector Applicability** data: strict consensus of 3334 most parsimonious trees retained after RC-based Character Reweight.

Fig. 8. Phylogenetic tree retained using only **Single Scalars** data: strict consensus of 24 most parsimonious trees.

Fig. 9. Phylogenetic tree retained using only **Single Scalars** data: the most parsimonious tree retained after triple RC-based Character Reweight.

Fig. 10. Phylogenetic tree retained using only **Single Angles** data: strict consensus of 1000 most parsimonious trees.

Fig. 11. Phylogenetic tree retained using only **Single Angles** data: strict consensus of 1000 most parsimonious trees after RC-based Character Reweight.

Fig. 12. Phylogenetic tree retained using only **Scalar Ratios** data: strict consensus of 2 most parsimonious trees.

Fig. 13. Phylogenetic tree retained using only **Scalar Ratios** data: the most parsimonious tree retained after triple RC-based Character Reweight.

Fig. 14. Phylogenetic tree retained using only **Pairwise Angles** data: the most parsimonious tree.

Fig. 15. Phylogenetic tree retained using only **Pairwise Angles** data: the most parsimonious tree retained after triple RC-based Character Reweight.

Fig. 16. Phylogenetic tree retained using only **Vector Differences** data: strict consensus of 2 most parsimonious trees.

Fig. 17. Phylogenetic tree retained using only **Vector Differences** data: the most parsimonious tree retained after triple RC-based Character Reweight.

Fig. 18. Phylogenetic tree retained using **all data types**: the most parsimonious tree.

Fig. 19. Phylogenetic tree retained using **all data types**: the most parsimonious tree retained after triple RC-based Character Reweight.

Fig. 20. Phylogenetic tree retained using **Vector Applicability + Scalar Ratios** data: the most parsimonious tree.

Fig. 21. Phylogenetic tree retained using **Vector Applicability + Single Scalars + Scalar Ratios** data: strict consensus of 2 most parsimonious trees.

Fig. 22. Phylogenetic tree retained using **Vector Applicability + Scalar Ratios + Vector Differences** data: the most parsimonious tree.

Fig. 23. Phylogenetic tree retained using **Vector Applicability + Scalar Ratios**: the most parsimonious tree retained after triple RC-based Character Reweight.

Fig. 24. Phylogenetic tree retained using **Vector Applicability + Single Scalars + Scalar Ratios**: the most parsimonious tree retained after triple RC-based Character Reweight. Bootstrap supports above 50% are shown.

Fig. 25. Phylogenetic tree retained using **Vector Applicability + Scalar Ratios + Vector Differences**: the most parsimonious tree retained after triple RC-based Character Reweight.

Tables

Table 1. The list of Cretaceous Ichneumonidae, with references and color mapping used in the analysis

* Poorly preserved specimens not included into the analysis

** New genera and species of Novichneumoninae from Burmese amber, not formally described yet.

*** *Heteropimpla megista*, described in Labenopimplinae, obviously should be transferred to Novichneumoninae

Table 2. The list of landmarks used in the analysis.

Table 3. The list of vectors used in the analysis.

Supplementary data

Supplement 1. Program VGMA v.0.2.2 alpha. For Windows only. No installation needed.

Supplement 2. Source code for VGMA v.0.2.2 alpha, Visual Studio C# project.

Supplement 3. Quick guide for VGMA. English and Russian versions.

Supplement 4. Source data for the case study analysis: images and VGMA map file.

Supplement 5. The results of phylogenetic analyses: Nexus files, PAUP logs, retained trees.

Figures

Figure 1

The example of Vector Comparison visualization of some Cretaceous ichneumonid fore wings: A – *Tanychora petiolata*, B – *Palaeoichneumon freja*, C – *Novichneumon longus*. Red lines correspond to veins shorter, than average, and green ones – longer, than average.

Figure 2

Pairwise Scalar visualization of comparative $2+3-M$ and $1-Cu$ lengths in Cretaceous ichneumonids.

Figure 3

Pairwise Angle visualization of the angle between 2-Rs and 2+3-M in Cretaceous ichneumonids. Solid lines represent the spread of values within subfamilies, sectors represent the spread excluding outliers.

Figure 4

Pairwise Vector visualization of vector subtraction between pterostigma length and 2-Rs.

0 - MAX

MIN - MAX

Log MIN - MAX

Figure 5

Methods of conversion of morphometric data into phylogenetic code implemented in VGMA.

Figure 6

Phylogenetic tree retained using only Vector Applicability data: strict consensus of 46045 most parsimonious trees. Here and in other trees Tanychorinae are marked with red dots, Palaeoichneumoninae – with green, Labenopimplinae – with blue, Novichneumoninae – with pink, Labeninae – with orange and Tryphopimpla – with grey.

Figure 7

Phylogenetic tree retained using only Vector Applicability data: strict consensus of 3334 most parsimonious trees retained after RC-based Character Reweight.

Figure 8

Phylogenetic tree retained using only Single Scalars data: strict consensus of 24 most parsimonious trees.

Figure 9

Phylogenetic tree retained using only Single Scalars data: the most parsimonious tree retained after triple RC-based Character Reweight.

Figure 10

Phylogenetic tree retained using only Single Angles data: strict consensus of 1000 most parsimonious trees.

Figure 11

Phylogenetic tree retained using only Single Angles data: strict consensus of 1000 most parsimonious trees after RC-based Character Reweight.

Figure 12

Phylogenetic tree retained using only Scalar Ratios data: strict consensus of 2 most parsimonious trees.

Phylogenetic tree retained using only Scalar Ratios data: the most parsimonious tree retained after triple RC-based Character Reweight.

Figure 14

Phylogenetic tree retained using only Pairwise Angles data: the most parsimonious tree.

Figure 15

Phylogenetic tree retained using only Pairwise Angles data: the most parsimonious tree retained after triple RC-based Character Reweight.

Figure 16

Phylogenetic tree retained using only Vector Differences data: strict consensus of 2 most parsimonious trees.

Figure 17

Phylogenetic tree retained using only Vector Differences data: the most parsimonious tree retained after triple RC-based Character Reweight.

Figure 18

Phylogenetic tree retained using all data types: the most parsimonious tree.

Figure 19

Phylogenetic tree retained using all data types: the most parsimonious tree retained after triple RC-based Character Reweight.

Figure 20

Phylogenetic tree retained using Vector Applicability + Scalar Ratios data: the most parsimonious tree.

Figure 21

Phylogenetic tree retained using Vector Applicability + Single Scalars + Scalar Ratios data: strict consensus of 2 most parsimonious trees.

Figure 22

Phylogenetic tree retained using Vector Applicability + Scalar Ratios + Vector Differences data: the most parsimonious tree.

Figure 23

Phylogenetic tree retained using Vector Applicability + Scalar Ratios: the most parsimonious tree retained after triple RC-based Character Reweight.

Figure 24

Phylogenetic tree retained using Vector Applicability + Single Scalars + Scalar Ratios: the most parsimonious tree retained after triple RC-based Character Reweight. Bootstrap supports above 50% are shown.

Figure 25

Phylogenetic tree retained using Vector Applicability + Scalar Ratios + Vector Differences: the most parsimonious tree retained after triple RC-based Character Reweight.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Suppl.1.VGMAv.0.2.2a.zip](#)
- [Suppl.2.Sourcecode.zip](#)
- [Suppl.3.Quickguide.zip](#)
- [Suppl.4.Sourcedata.zip](#)
- [Suppl.5.Phylogeneticanalysis.zip](#)
- [Table1.ListofCretaceousIchneumonidae.pdf](#)
- [Table2.Listoflandmarks.pdf](#)
- [Table3.Listofvectors.pdf](#)