APPLYING GEOMETRIC MORPHOMETRICS TO THE CLASSIFICATION AND IDENTIFICATION OF SMALL MAMMALS – A Case Study of Large Samples of Four Arvicoline Species

Article · October 2012

5 authors, including:

Vertebrata Palasiatica

Chinese Academy of Sciences

213 PUBLICATIONS   **461** CITATIONS

古 脊 椎 动 物 学 报
VERTEBRATA PALASIATICA

几何形态测量学方法在小哺乳动物化石分类鉴定中的应用[1)]
—— 4种鼩类化石大样本的个案研究

张颖奇[1]　严亚玲[1]　刘毅弘[1]　魏光飚[2]

(1 中国科学院古脊椎动物与古人类研究所，脊椎动物进化系统学重点实验室　北京　100044)

(2　重庆中国三峡博物馆三峡古人类研究所　重庆　400015)

摘要： 小哺乳动物化石在晚新生代生物地层学和生物年代学研究中具有重要作用，尤其是筛洗法在古生物调查中得到广泛应用后，通常可以采集到数量可观的标本，使得其地位较大型哺乳动物化石更加彰显。因此，小哺乳动物化石标本的分类鉴定也成了一项十分关键的工作。然而，传统形态学方法在对大量标本进行分类鉴定时，往往容易受主观因素影响而将不稳定的细微性状变异作为依据建立新种，或者忽视一些肉眼难于察觉的形态学差异而将两个甚至多个类群合并到一起，导致基于形态学的化石分类鉴定随意性增加，失去客观性。此外，对于不具鉴定意义的非关键性单个牙齿，很难凭借肉眼或显微镜观察进行区分。针对这些问题，本文选取了安徽繁昌人字洞早更新世早期三种鼩类甘肃模鼠*Mimomys gansunicus*，郑氏异模鼠*Heteromimomys zhengi*和繁昌维蓝尼鼠*Villanyia fanchangensis*的1284件臼齿并以早上新世内蒙古比例克的比例克模鼠*Mimomys bilikeensis*的163件臼齿作为参考，采用几何形态测量学方法在各个臼齿咬合面上分别选取了7~14个同源landmark对咬合面的形态特征进行了线性判别分析，建立了针对这4个种的臼齿咬合面形态差异判别函数。分析结果表明，根据这些landmark所提供的形态差异信息，人字洞的1284件臼齿标本中的确存在3个可以明显区分且形态学性状稳定的类群，先前的分类鉴定得到了验证。与之不同地点不同时代的比例克模鼠也可以很好区分。因此，建立在大样本基础上的这4种鼩类臼齿咬合面同源形态特征的判别函数可以用来描述这些种类较为稳定的形态特征差异，并用来作为今后对样本较少的标本进行分类鉴定时的判别依据。因为几何形态测量学的方法不仅可以对二维的离散landmark数据以及重要形态学特征的轮廓线进行分析，甚至还可以扩展至三维空间，所以上述方法对于较为容易获得大样本标本的小哺乳物化石分类鉴定具有普遍意义。

关键词： 几何形态测量学，线性判别分析，小哺乳动物，分类鉴定

APPLYING GEOMETRIC MORPHOMETRICS TO THE CLASSIFICATION AND IDENTIFICATION OF SMALL MAMMALS
—A Case Study of Large Samples of Four Arvicoline Species

ZHANG Ying-Qi[1]   YAN Ya-Ling[1]   LIU Yi-Hong[1]   WEI Guang-Biao[2]

(1 *Key Laboratory of Evolutionary Systematics of Vertebrates, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences* Beijing 100044   arvicolid@gmail.com)

(2 *Chongqing Three Gorges Institute of Paleoanthropology, China Three Gorges Museum* Chongqing 400015)

Abstract   Small mammals play a very important role in late Cenozoic biostratigraphy and biochronology. Especially since screen-washing technology began to be applied in paleontological investigations, it has often been possible to acquire large samples of small mammal fossils, which makes them a more effective proxy for biostratigraphy and biochronology than large mammals. Accordingly, classification and identification of these large samples have become a fundamental paleontological task. However, when dealing with them using empirical or intuitive traditional morphological methods, it is impossible to avoid paying too much attention to some slight and inconsistent differences and "oversplitting" taxonomic units, or conversely neglecting some consistent differences that are imperceptible to human eyes and lumping several taxonomic units together. In either case, subjective factors will introduce a degree of arbitrariness into the classification and identification of the specimens. To avoid this, a quantitative geometric morphometric method of classification and identification based on analysis of large samples is explored in this paper. The subjects of this case study are 1284 isolated molars of *Mimomys gansunicus*, *Heteromimomys zhengi* and *Villanyia fanchangensis* from the lower Lower Pleistocene Renzidong site, and 163 specimens of *Mimomys bilikeensis* from the Lower Pliocene Bilike site are used as a point of reference. 7-14 2D homologous landmarks were defined on the occlusal surface of each of the six molars and used as the basis for a linear discriminant analysis. The results confirm that there are three different arvicoline species in the Renzidong sample, and the linear discriminant functions produced in the analysis can describe the consistent differences that exist among the species in this large sample. Furthermore, the same functions can be used as a basis for identifying newly recovered fossil specimens of related arvicoline species. Because geometric morphometrics can deal with both discrete landmark data and continuous outline data pertaining to significant morphological characters, and is suitable for use in both 2D and 3D, this method can be generally applied to the classification and identification of small mammal fossils.

Key words   geometric morphometrics, linear discriminant analysis, small mammals, classification and identification

1 Introduction

Taxonomic practice is still characterized by an impasse between "lumpers" and "splitters". Taxonomists belonging to either category are usually thought of as being limited in their objectivity, because they base their classifications and identifications on intuitive or subjective judgments. In practice, we are particularly likely to be faced with this problem when dealing with small mammal fossils, because use of the screen-washing technique in paleontological surveys of terrestrial deposits often makes it possible to acquire large samples of such fossils. The greater abundance of small mammal as opposed to large mammal fossils is one of the reasons why small mammals take priority in late Cenozoic biostratigraphy and biochronology. However, large sample sizes also lead to the problem of subjectivity when we classify and identify the fossils. On one hand, there is a strong possibility that too much attention becomes focused on slight differences among specimens. Attaching too much weight to such variations leads to excessive "splitting" of taxonomic units. On the other hand, some actual consistent differences are too slight to be recognizable to the human eye, and failure to perceive these differences can cause different taxa to be "lumped" together. Therefore, objective criteria for recognizing interspecific boundaries are needed in order to overcome the subjective disagreements between "lumpers" and "splitters" that can arise when coping with large samples of small mammal fossils.

In contrast to conventional qualitative approaches that rely mainly on intuition and experience, geometric morphometrics is a statistics-based quantitative way of comparing shape (morphology) across different specimens. It can detect similarities and differences, even slight ones, among the individuals included in the analysis. Geometric morphometrics not only offers the ability to describe shape precisely and accurately, but also facilitates visualization and interpretation of results of the analysis. Because the method is statistics-based, a large sample size is required in order to obtain statistically significant results. Of course, a larger sample size can be expected to lead to more robust and reliable results.

The objective of this paper is to demonstrate how to apply geometric morphometric method to the classifications and identifications of small mammals to avoid subjectivity based on a case study of large sample.

2   Material and methods

2.1   Material

The early Early Pleistocene Renzidong site, in Anhui Province, has yielded more than 2500 arvicoline specimens, referable to the three species *Heteromimomys zhengi*, *Villanyia fanchangensis*, and *Mimomys gansunicus* (Zhang et al., 2008, 2010; Jin and Liu, 2009).

The Early Pliocene Bilike site in Nei Mongol has also yielded more than 2000 arvicoline specimens, all of which have been referred to the single species *Mimomys* (*Aratomys*) *bilikeensis* (Qiu and Storch, 2000; Repenning, 2003). The specimens from these localities constitute a large sample of teeth that each carries an *a priori* attribution to one of four arvicoline species. In order to produce robust and reliable results, the numerous teeth comprising this sample were selected as the subjects of the geometric morphometric analysis in this study (Table 1). Arvicoline molars have hypsodont, triangularly prismatic cusps. As a result, the morphology of the occlusal surfaces of the molars remains relatively constant during most of the wear process, permitting occlusal surface morphology to act as one of the main criteria in the classification and identification of fossil arvicolines. None of the teeth selected for this study have suffered more than light or moderate wear. The *Mimomys bilikeensis* sample is used as a reference to testify whether the analysis performed can detect the actual consistent differences.

Table 1   Number of available specimens, and landmarks, for each type of element analyzed in this study

Locality	Species	m1	m2	m3	M1	M2	M3
	Heteromimomys zhengi	50	21	33	28	30	26
Renzidong	*Villanyia fanchangensis*	406	38	36	52	47	187
	Mimomys gansunicus	147	29	25	33	32	64
Bilike	*Mimomys bilikeensis*	49	21	18	25	24	26
	Total	652	109	112	138	133	303
	Landmark number	14	11	11	11	9	7

2.2   Landmark selection

Landmarks play a fundamental role in geometric morphometrics. Ideally, each landmark should be homologous across all specimens included in the analysis. From 7 to 14 landmarks were established on the occlusal surface of each of the six molar types included in the study (Table 1; Fig.1). Basically, all the vertices of the reentrant and salient angles on each molar were designated as landmarks, because they can be accurately located and can obviously be thought of as homologous. The anteriormost points on m1 and m2, the posteriormost points on M1 and M2, were also selected. Among the four species analyzed, only *Heteromimomys zhengi* has a very strongly developed so-called *Mimomys*-angle. This structure is very weak in *Mimomys bilikeensis*, and has a very low frequency of occurrence in the *Mimomys gansunicus* sample. When present in *M. gansunicus*, it is vestigial and not as close BSA3 (buccal salient angle 3) as it is in *H. zhengi* and *M. bilikeensis*, which suggests that landmark 13 on m1 might fail to meet the criterion of homology. Is the vestigial *Mimomys*-angle in *M. gansunicus* homologous with that in *H. zhengi* and *M. bilikeensis* or not? Here we choose the latter, taking into account that the vestigial *Mimomys*-angle probably represents a novel salient angle in development. So the landmark 13 is finally positioned on the vertex of BRA3 (buccal reentrant angle 3) in all four species. Another problem is that, because *M. bilikeensis* represents a very

primitive form of the genus, BRA3 is sometimes too flat to be noticeable in the sample. In such situations, landmark 13 is positioned on the first vertex of the enamel band curve immediately in front of MA (the *Mimomys*-angle), or immediately in front of BSA3 when MA cannot be observed. All landmarks used in this study are type II according to the classification of Bookstein (1991).

2.3 Landmark data acquisition and analysis

R is a language and environment for statistical computing and graphics (R Development Core Team, 2010). It can be adapted to perform any morphometric analysis (Claude, 2008). The whole process of geometric morphometric analysis, from landmark data acquisition to final graphing of results, can be performed in the R command-line based environment. Based on this factor, R was chosen over other geometric morphometric analysis software for the present study, in order to benefit from the convenience of using a single program. The command scripts needed to carry out the entire of analysis were mainly modified from Claude (2008) by the authors.

2.3.1 Landmark data acquisition

Photos of the occlusal surfaces of all specimens in the samples included in the analysis were taken using an OMRON 3CCD digital fine scope system (VC4500-PC). Each specimen was positioned with the occlusal surface parallel to the camera lens, and centered in the effective field of view of the camera. Before the photos were read into R one by one, all photos of right teeth were reversed to achieve consistency with left teeth, and then the landmark configurations are collected according to Fig.1. The read.jpeg() function of the "rimage" package, and the locator() function of the "graphics" package, were mainly utilized in these steps.

2.3.2 Procrustes superimposition of landmarks

Procrustes superimposition of landmarks was carried out in order to remove shape-irrelevant variables like size, orientation and position from the original landmark configurations, leaving the real shape information. This is a necessary step in geometric morphometric analysis. The functions aligne(), angle2d(), angle3(), centsiz(), mshape(), orp(), pgpa(), pPsup(), and trans1() developed by Claude (2008) were utilized. The superimposed landmarks were then used as parameters in the subsequent linear discriminant analysis.

2.3.3 Linear discriminant analysis (canonical variate analysis)

The purpose of linear discriminant analysis (or canonical variate analysis for more than two groups) is to quantitatively describe inter-group differences and predict the attribution of a new observation by the discriminant function. It requires *a priori* categorization of the subjects of the analysis. In this case study, the subjects were categorized *a priori* into four
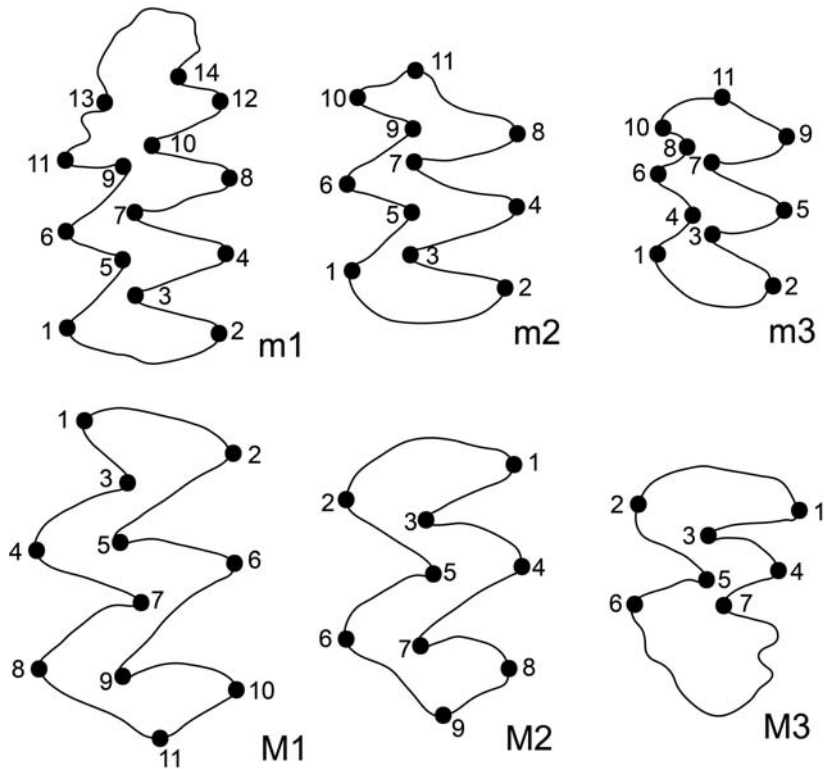
Fig. 1   Landmarks used in this study

species, and the analysis was intended to determine whether or not this initial classification and identification of the specimens was quantitatively acceptable. Accordingly, there were three discriminant axes. However, because linear discriminant analysis cannot test the statistical significance of the intergroup differences it identifies, multivariate analysis of variance was also utilized. Specifically, we employed Wilks' lambda test to determine if the linear discriminant functions were statistically significant. LDs (linear discriminant functions or axes) that passed the test could be used as effective discriminators to predict which group (i.e. species) a newly discovered tooth would be attributed to. The main functions used for this step were lda() in the "MASS" package, predict() in the "stats" package and Wilks.test() in the "rrcov" package.

3   Results

The results of the linear discriminant analysis (LDA) and Wilks' lambda test on the LDs are briefly summarized in Table 2. The "proportion of trace" values given for the LDA represent the percentage of total between-group variation that can be explained by each LD. The NULL hypothesis of the Wilks' lambda test is equality of group means. The extremely low p-values obtained for each of the six molars indicate that the NULL hypothesis can be rejected, and therefore that the differences explained by the three LDs are statistically significant.

Table 2 LDA and Wilks' lambda test results

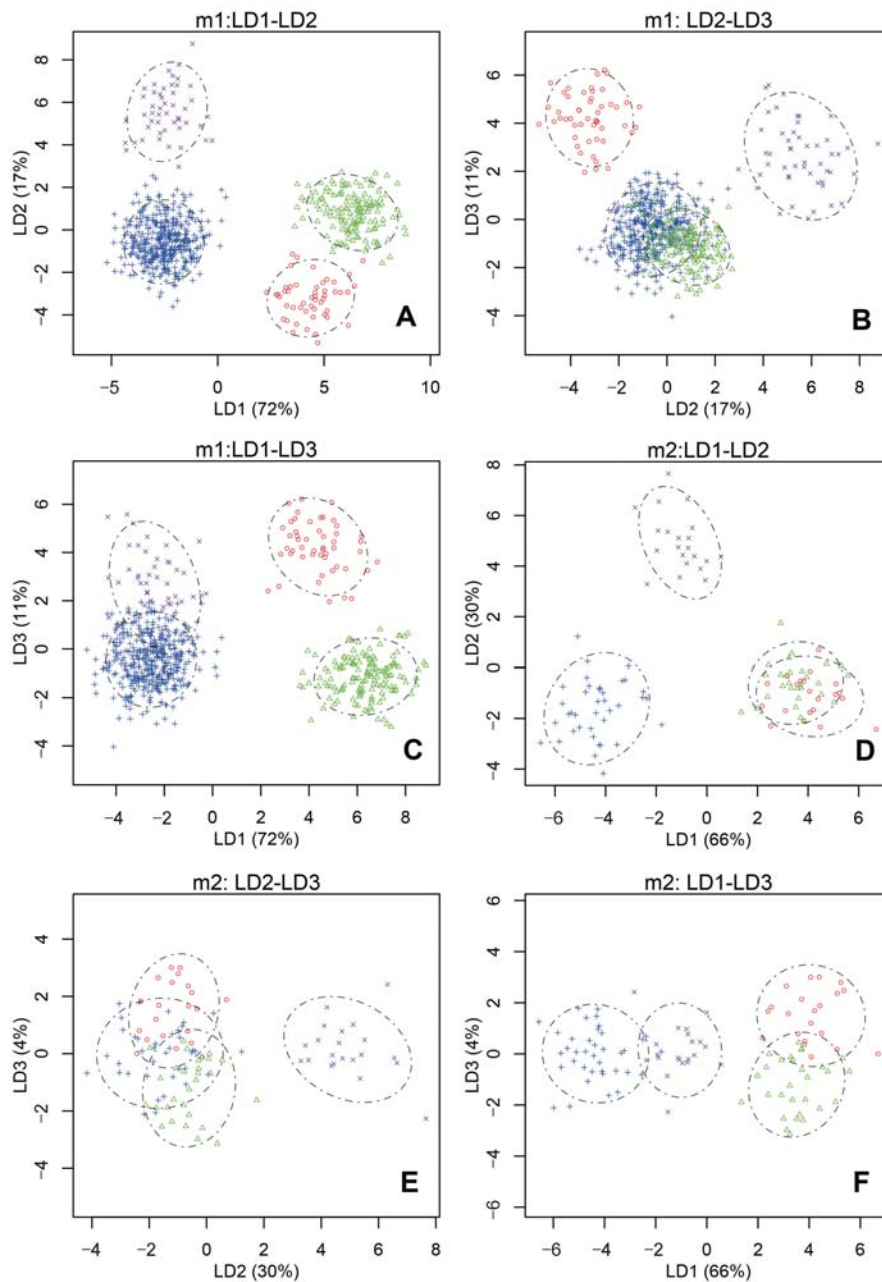|	LDA: Proportion of trace			Wilks' lambda test			
	LD1	LD2	LD3	Wilks' lambda	Chi2-value	DF	p-value
m1	0.7221	0.1670	0.1109	0.0042	3547.071	9	< 2.2e−16
m2	0.6626	0.2965	0.0410	0.0051	551.835	9	< 2.2e−16
m3	0.5345	0.3765	0.0890	0.0057	554.685	9	< 2.2e−16
M1	0.6749	0.2972	0.0279	0.0101	613.903	9	< 2.2e−16
M2	0.5494	0.3623	0.0883	0.0104	587.028	9	< 2.2e−16
M3	0.6027	0.2973	0.0999	0.0549	866.109	9	< 2.2e−16

Fig. 2 Biplots of the LDA results for m1 (A—C) and m2 (D—F)

Red ○: *Heteromimomys zhengi*; green △: *Mimomys gansunicus*; blue +: *Villanyia fanchangensis*; purple ×: *Mimomys bilikeensis* (the same for the following figures)

Biplots with 95% confidence ellipses for the three LDs of each of the six molars are shown in Figs.2—4. These show the positions of all the individual molars in LD1-LD2-LD3 space. Only for m1 and m3 are the 95% confidence intervals of all four species completely separate. By contrast, the results for m2, M1 and M2 show varying degrees of overlap between the confidence intervals for *Mimomys gansunicus* and *Heteromimomys zhengi*. The M3 shows slight overlap among the three species from Renzidong, while *Mimomys bilikeensis* is totally separate from all three. Despite the varying degrees of overlap, the Wilks' lambda test indicates that the differences among the four species are statistically significant. In other words, the four
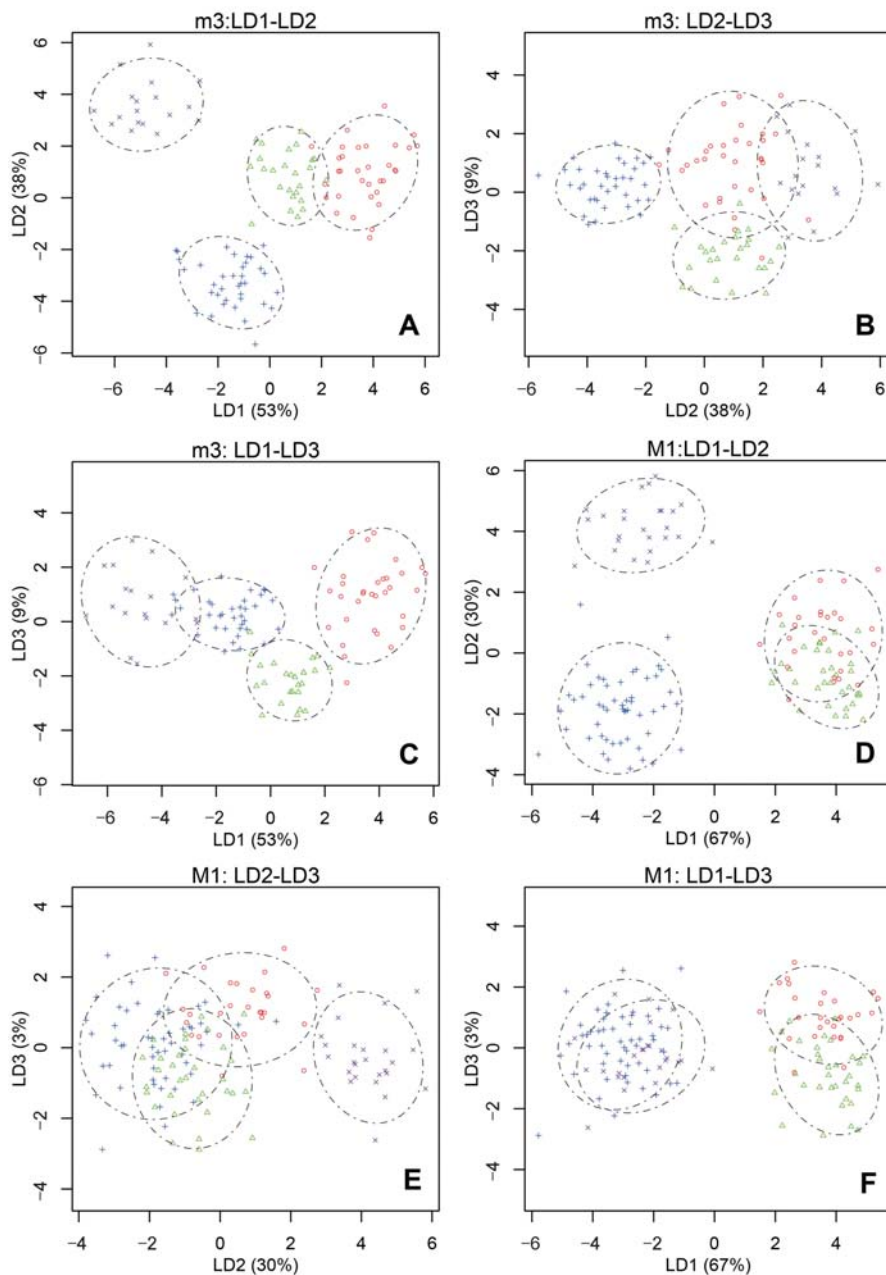

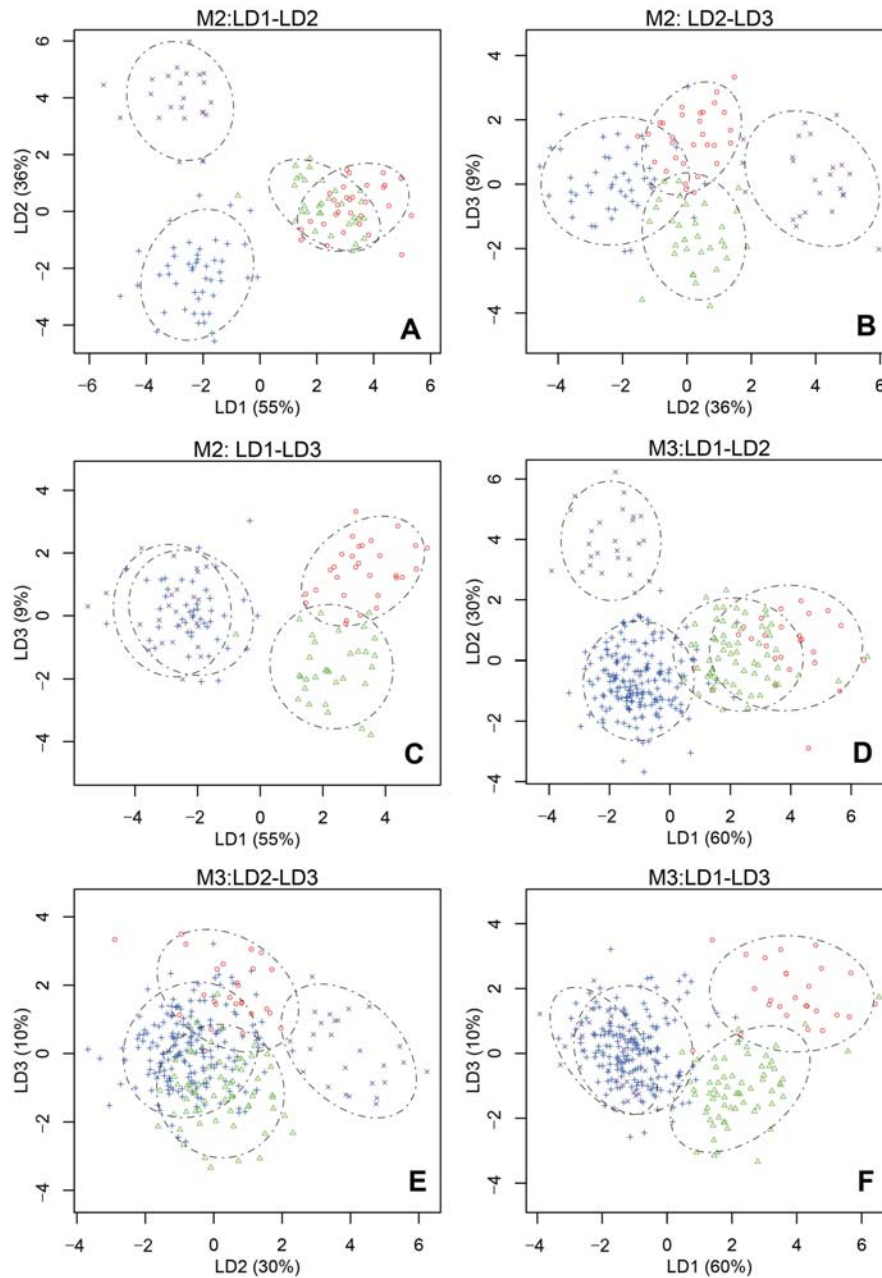Fig. 3   Biplots of the LDA results for m3 (A—C) and M1 (D—F)

Fig. 4   Biplots of the LDA results for M2 (A—C) and M3 (D—F)

species differ significantly from one another in dental morphology and the LDs obtained in the analysis above can be used as a stable basis for discriminating among them.

4   Discussion

4.1   *A priori* grouping of samples

Linear discriminant analysis requires *a priori* grouping of the objects being analyzed in order to generate reliable linear discriminant functions. In our case study, this was

accomplished by adopting the previously published classifications and identifications (Zhang et al., 2008, 2010; Jin and Liu, 2009). This earlier published work does not conflict with the LDA results presented above. On the contrary, the latter confirms the presence of three stable arvicoline species in the Renzidong fauna. However, no ready-made groupings are available when dealing with untouched large samples of newly screen-washed small mammal fossils. In fact, there exists a more quantitative and objective alternative to the adoption of published classifications and identifications, namely the method of principal component analysis (PCA). PCA can transform the landmark data to a new coordinate system, which is defined so that the greatest amount of variance in the data is expressed by the first transformed new variable (called the first principal component), the second greatest amount of variance by the second new variable, and so on. In contrast to LDA, which focuses on differences among group means, PCA is used to describe differences among individuals. Furthermore, PCA does not require *a priori* grouping of samples. Applying PCA to newly recovered large samples of small fossils should help to determine, in quantitative term, how many morphologically different groups are present. The PCA results can thus provide a basis for the *a priori* grouping required by LDA in order to generate reliable linear discriminant functions. Fig.5 shows PCA results, including the first three principal components and measurements, for the m1 specimens included in the case study described in this paper. The results show that all four species can be easily distinguished
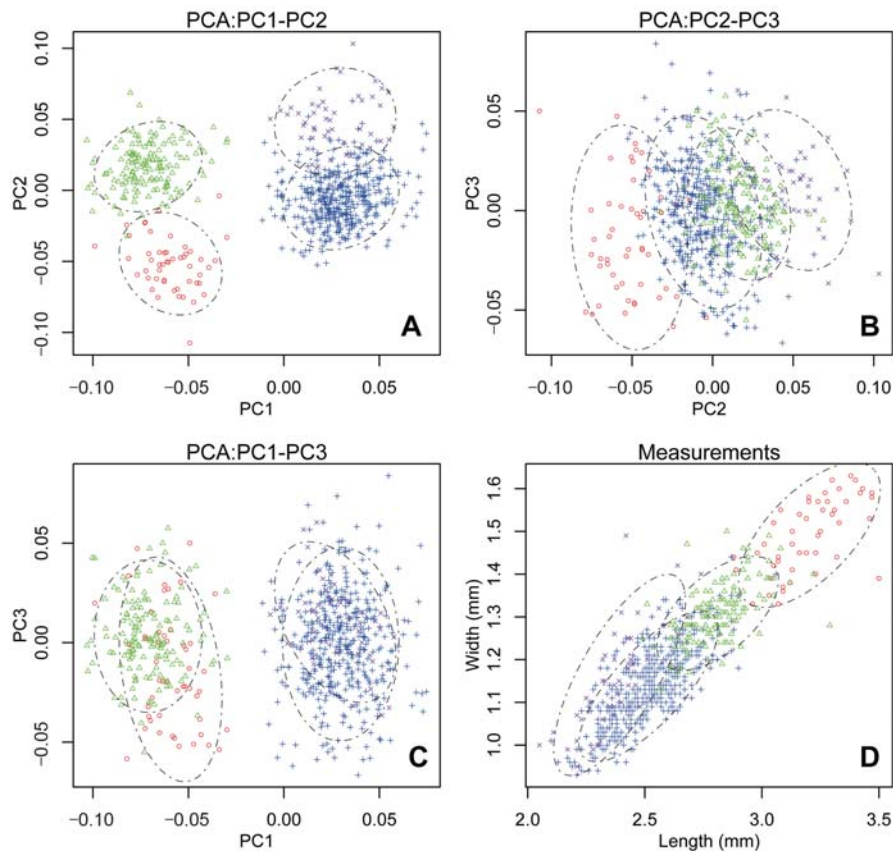

Fig. 5   Biplots of the PCA results (A-C) and measurements (D) for m1

from one another in the PC-PC2 space, apart from overlap between *Mimomys bilikeensis* and *Villanyia fanchangensis*. In fact, *M. bilikeensis* and *V. fanchangensis* overlap each other on all three PCA biplots and also the length and width measurements. This indicates a genuine resemblance in m1 morphology between these two species based on the landmarks established in this study, and aslo suggests that they may be closely related phylogenetically. Usually this will happen in practice. In such cases, one possible course of action is to accept the evidence of similarity between the samples, and treat them as a single taxonomic unit. Alternatively, if there are strong reasons to believe that the samples are in fact distinct, one should pick out the doubtful specimens and analyze them using another set of landmarks. They will be finally separated, or, on the contrary, proved to be the same species. In the present example, *M. bilikeensis* and *V. fanchangensis* are very different species that come from different localities and are of different geological ages. It is likely that the two species would be separated by a PCA using a different set of landmarks. Once the step of separation, or so-called "*a priori* grouping", is complete, and then you can "split" or "lump" them, and build up discriminant functions for new specimens.

4.2　Discrimination of new fossil samples

　　Once the linear discriminant functions for a given set of groups (species) have been obtained and shown to pass the Wilks' lambda test, they can be used to assign new samples to these same groups. The functions lda() in the "MASS" package and predict() in the "stats" package are used to achieve this purpose in the R environment. The following scripts and outputs demonstrate a successful example of assigning each of four randomly selected specimens from the samples used in this study to one of the four species. For a new observation to be discriminated, the whole process would be the same.

```
> mod1 < −lda (m,fact)
Warning message:
In lda.default (x, grouping, ...): variables are collinear
> predict (mod1, m[c(48,399,500,640),])
$class
[1] z f g b
Levels: b f g z

$posterior
              b              f              g              z
[1,] 7.712575e−18   2.096286e−12   4.883191e−09   1.000000e+00
[2,] 5.446266e−21   1.000000e+00   9.064608e−19   4.310352e−15
[3,] 4.877826e−28   1.040233e−16   1.000000e+00   1.758871e−11
[4,] 9.998233e−01   1.767307e−04   1.984516e−19   1.487758e−26

$x
         LD1         LD2         LD3
[1,]  3.518100   −1.1456510    4.6785806
[2,] −2.114597   −3.6196384   −1.2280443
[3,]  6.117739   −0.1787128   −1.8183982
[4,] −2.494714    4.6833168    0.2348492
```

The underlined numbers represent the probability that the specimen in question belongs to each of the four species (b: *Mimomys bilikeensis*; f: *Villanyia fanchangensis*; g: *Mimomys gansunicus*; z: *Heteromimomys zhengi*). The double-underlined numbers represent the highest probabilities calculated for the individual specimens, which also indicate their final taxonomic assignments as determined by this analysis. The correct assignments for the four specimens should be z, f, g, b, in order, so the analysis has identified them correctly.

4.3　Data collection for geometric morphometric analysis

Data collection is fundamentally important in geometric morphometric analysis. There are two types of geometric morphometric data, namely landmark data and outline data. With R, both types of data can be handled and analyzed both in 2D and 3D. In this case study, the subjects are arvicoline molars from various species. Because the occlusal surfaces of the molars are fairly flat, 2D landmark data can be quite easily acquired by photographing them. For small mammal taxa in which the teeth have more three-dimensional morphology, such as murids or insectivores, 3D landmark data or outline data may be too difficult to acquire, but 2D data can still be collected by photographing the teeth in appropriate views. No matter what type of data are collected, however, they must be morphologically and taxonomically meaningful in order to lead to positive results.

References

Bookstein F L, 1991. Morphometric Tools for Landmark Data: Geometry and Biology. Cambridge: Cambridge University Press. 1−456

Claude J, 2008. Morphometrics with R (Use R). New York: Springer Science+Business Media, LLC. 1−316

Jin C Z(金昌柱), Liu J Y(刘金毅), 2009. Paleolithic Site—The Renzidong Cave, Fanchang, Anhui Province. Beijing: Science Press. 1−439(in Chinese with English summary)

Qiu Z D, Storch G, 2000. The Early Pliocene micromammalian fauna of Bilike, Inner Mongolia, China (Mammalia: Lipotyphla, Chiroptera, Rodentia, Lagomorpha). Senckenbergiana lethaea, **80**(1): 173−229

R Development Core Team, 2010. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. ISBN 3-900051-07-0, URL http://www.R-project.org/

Repenning C A, 2003. *Mimomys* in North America. Bull Am Mus Nat Hist, **279**: 469−512

Zhang Y Q, Kawamura Y, Jin C Z, 2008. A new species of the extinct vole *Villanyia* from Renzidong Cave, Anhui, East China, with discussion on related species from China and Transbaikalia. Quatern Int, **179**: 163−170

Zhang Y Q, Jin C Z, Kawamura Y, 2010. A distinct large vole lineage from the Late Pliocene-Early Pleistocene of China. Geobios, **43**(4): 479−490