Morphological Variations and Discriminant Analysis of Two Populations of *Coilia ectenes* *

CHENG Qiqun[1,2] & HAN Jindi[1]

(1: Key Laboratory of Marine and Estarine Fisheries, Ministry of Agriculture, East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai 200090, P. R. China;
2: Institute of Genetics, Fudan University, Shanghai 200433, P. R. China)

Abstract Traditional morphological data and truss network data were combined to conduct principal component analysis and stepwise discriminant analysis to study the morphological variations between two populations of *Coilia ectenes*. The results are showed as follows: morphological variations between these two populations were mainly caused by the distance from most anterior of scales on skull to origin of dorsal fin; discriminant analysis with 25 morphological parameters indicated that there were extremely significant differences between the two populations of Coilia ectenes ($P < 0.01$), the accuracy is 100%; discriminant analysis with selected 9 morphological parameters showed that there were extremely significant differences between the two populations of *Coilia ectenes* ($P < 0.01$), the identification accuracy are 95% - 100% (P1) and 95.2% - 100% (P2) respectively, and the synthetic identification accuracy is 97.5%; ANOVA analysis was conducted on the morphological parameters of *Coilia ectenes* and the coefficients of difference of these parameters were also calculated. According to Mayr's 75% Rule, this paper reveals that the difference between the two populations of *Coilia ectenes* belongs to population level, and has not risen to the subspecies level.

Keywords *Coilia ectenes*, lake anchovy (*coilia ectenes*), Discriminant analysis, Principal Capital Analysis, Truss analysis, Multivariation analysis

1 Introduction

Coilia ectenes, belongs to *Coilia*, Engraulidae, Clupeiformes, and it is one of the most important economic fishes in the estuary of Yangtze River, P. R. China. It is an anadromous fish. Matured fish make productive migration in succession from near ocean to river when reproductive season comes. Matured fish migrate against the river and spawn respectively in the lower reaches and middle reaches of *Yangtze* River, and also they spawn in the lakes that connect Yangtze River[1]. Most of the matured fish migrate back to near ocean and live through the winter after they have spawned, but still some matured fish settle down in the lakes that pertain to the

* Supported by the Youth Foundation of Chinese Academy of Fisheries Sciences (2003 - youth - 4).
Received: 2004 - 02 - 23; Accepted: 2004 - 06 - 16. Cheng Qiqun, male, born in 1972, assistant professor, phD candidate, E-mail : qiquncheng@ yahoo. com. cn.

Yangtze River, and they acquire some distinctive differences with those who migrate regularly in ecological characters and breeding habits with the changing of environmental factors, and form another ecological group who have smaller body. Lake anchovy (*Coilia ectenes*), which is an e-cological group of *Coilia ectenes* who migrate into Taihu lake and settle down there, so it is an ecesis group of *Coilia ectenes*.

From ecological characters, we know see that Lake anchovy (*Coilia ectenes*) is earlier than *Coilia ectenes* in the age of sex maturity, shorter in body length of matured fish, shorter in matured egg diameter, later in the booming period of spawn, and *Coilia ectenes* have plumper liver than Lake anchovy (*Coilia ectenes*) [1]. All these characters show that there are prodigious differences between these two populations.

In fisheries sciences, there are no reports on the morphological differences between migrating group and ecesis group till now. So it is an interesting thing to disclose the morphological differences between them when separated for long time. To elucidate this question, in this study, we conduct principal component analysis and stepwise discriminant analysis to study the morphological variations between two populations of *Coilia ectenes*, and establish corresponding discriminant formula; also ANOVA analysis were conducted with every morphological character parameters of them, then coefficient of difference (CD) were calculated and their morphological differences were classified according to Mayr's 75% Rule[2].

2 Materials and methods

2.1 Fish samples

The samples include two populations of *Coilia ectenes*, 40 individuals of each population were collected. *Coilia ectenes* were collected at Chongming island(N31.31°, E121.8°), Shanghai city, and Lake anchovy were collected at Taihu Lake, Wuxi City, Jiansu Provinee(N31.38°, E120.1°)in Dec. 2002.

2.2 Data collection

2080 morphological data of 80 fish, e.g., 26 morphological data of each fish, were collected.

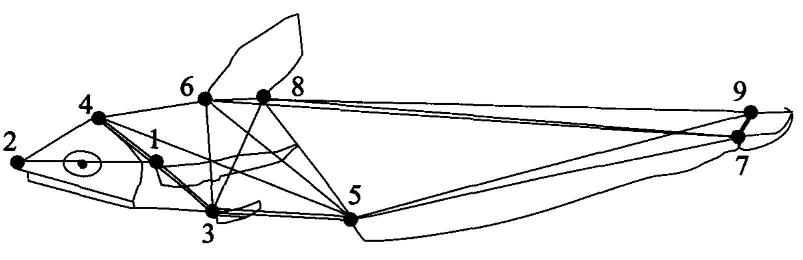

Fig. 1　Truss network of *Coilia ectenes* 18 truss parameter measurements are the distances between the two of 9 landmark points. For example, D1 - 2 denotes the distance between landmark points 1 and 2

Landmark points:

1. most posterior point of maxilla; 2. tip of snout; 3. origin of pelvic fin; 4. most anterior of scales on skull; 5. origin of anal fin; 6. origin of dorsal fin; 7. tremunus of anal fin; 8. tremunus of dorsal fin; 9. dorsal origin of caudal fin

They were constituted of traditional measureable characters and truss parameters. The selection of landmark points was chiefly based on Brzeski et al. [3] and Nancy et al. [4], only landmark point 2 at the head was added, so the information of former head was included. Truss network of *Coilia ectenes* was showed in Fig 1.

2.3 Data analysis

All 26 parameters, e. g., 8 measureable characters and 18 truss parameters, of each fish were treated first with Microsoft excel soft, then analysed with SAS software package. The methods were principal component analysis and stepwise discriminant analysis. To avoid the influence of body size difference on analysis, all parameters were corrected by dividing by their total length each. All parameters mentioned next were all corrected. Discriminant formula of two population was constructed based on those parameters which contribute more to their morphological difference. The coefficiency of difference of two population were also calculated.

Principal component analysis is a statistical method which transform many indexes into a few indexes. In this study, 4 independent synthetic indexes were calculated based on 26 parameters, e. g., PC1, PC2, PC3, and PC4. Contribute rate and accumulating contribute rate of principle components were calculated according to Zhang et al. [5].

Three accuracy of discriminat were calculated as follows:

(1) Accuracy of discriminant P1 = Number of discriminated accurately / Number of fact,

(2) Accuracy of discriminant P2 = Number of discriminated accurately / Number of discriminated,

(3) Accuracy of synthetic discriminant = $\sum_{i=1}^{n} Ai / \sum_{i=1}^{n} Bi$

in these formulas above, Number of fact means the number of this population collected, Number of discriminated means the number of discriminated into this population; in this study, Number of fact of these two population are both 40, Number of discriminated are *Coilia ectenes* 42 and lake anchovy (*coilia ectenes*) 38, respectively; Ai is number of discriminated accurately of the ith population, and Bi means the number of fact of the ith population, n is the number of population.

Coefficency of differences (CD) were calculated according to Mayr et al. [2].

$CD = (M_B - M_A) / (SD_A + SD_B)$,

in this formula, M_A and M_B are parameter means of population A and B separately, SD_A and SD_B are parameters standard deviation of population A and B separately. If CD of one parameter <1.28, then we can classify the differences between them into different geography population within species.

3 Results

3.1 Principal component analysis (PCA)

Contribuate rates of the first four pincipal components and the loadings of all index on them by pincipal component analysis (PCA) were showed in Tab. 1.

Tab. 1 The loadings of the first four principle components for 25 characters of the Coilia ectenes

Parameters	PC1 31.62%	PC2 19.78%	PC3 14.14%	PC4 8.93%
X1	-0.061	0.074	0.036	0.065
X2	0.005	0.015	0.024	0.006
X3	0.012	-0.001	0.023	-0.008
X4	-0.025	0.021	0.032	0.059
X5	-0.001	-0.016	0.004	0.006
X6	0.024	0.033	0.014	0.014
X7	-0.009	0.007	-0.006	0.018
X8	0.021	-0.002	0.002	-0.066
X9	-0.010	0.008	0.028	0.023
X10	0.001	0.052	0.024	-0.019
X11	-0.004	0.008	0.006	0.039
X12	0.047	-0.055	0.063	-0.028
X13	0.018	-0.014	0.024	-0.065
X14	0.082	-0.029	-0.036	-0.037
X15	0.019	-0.028	0.006	-0.041
X16	-0.158	0.820*	-0.425	-0.311
X17	0.011	0.022	0.033	0.017
X18	-0.130	0.194	0.081	0.124
X19	-0.126	0.348	0.164	0.839*
X20	-0.021	0.039	0.006	0.018
X21	0.018	-0.059	-0.083	-0.061
X22	0.951*	0.262	0.097	0.032
X23	-0.034	-0.052	-0.019	-0.196
X24	-0.144	0.277	0.870*	-0.338
X25	-0.013	0.030	0.013	-0.052

* marked loading> 0.700

We can see from Tab. 1 that the accumulating contribute rate of these four principle components has arrived at 74.47%, so it means that morphological differences between them can be summarized by 4 separately principle components.

We can see that PC1 is mostly affected by character X22, PC2 is mostly affected by character X16, PC3 is mostly affected by character X24, and PC4 is mostly affected by character X19. Thereinto, PC1 reflects the distance from most anterior of scales on skull to origin of dorsal fin, so it shows that morphological difference of these two populations is mainly affected by this distance, but PC2, PC3, PC4 are all reflect the back part of body, so it means that the back part of body has some influences on their morphological difference.

3.2 Discriminant analysis

7 measureable characters (except total length) and 18 truss parameters, total 25 morpholog-

ical parameters, are combined to do discriminant analysis. The accuracy of discriminant is 100%, and the effect of discriminant is extremely significant ($P < 0.01$).

To enhance the convenience and practicality of this formula, 9 parameters , e. g., X4,X18, *X5* , *X14*, *X20*, *X2*, *X1*, *X11*, *X15*, which contribute more to the difference of morphological were selected from these 25 parameters, to build a new discriminant formula. These 9 parameters were arranged in tab. 2 according to the F value.

The results of discriminant based on 9 selected morphological parameter were showed in Tab. 3, the effect of discriminant was extremely significant($P < 0.01$).

Tab. 2 Variables (ranged by *F* test values) with high contribution in discriminant analysis

Parameter	Corresponding morphological index	*F* value
X4	SL/TL	79.215
X18	D7 - 6/TL	27.770
X5	ED/TL	14.397
X14	D5 - 4/TL	11.799
X20	D7 - 9/TL	8.600
X2	BW/TL	5.064
X1	BL/TL	4.061
X11	D3 - 4/TL	3.083
X15	D5 - 6/TL	2.313

Tab. 3 Results of discriminant analysis of two kinds of Coilia ectenes (based on 9 selected parameters)

Actual population	Population discriminated		Accuracy of discriminant		Accuracy of synthetic discriminant(%)
	Coilia ectenes	Lake anchovy (coilia ectenes)	P1(%)	P2(%)	
Coilia ectenes (40)	40	0	100	95.2	
Lake anchovy (coilia ectenes) (40)	2	38	95	100	97.5
Total	42	38			
Percentage (%)	52.5	47.5			

Discriminant formulas of these two populations based on 9 parameters were as follows:

Coilia ectenes

$$Y1 = -5305 + 7937X1 + 2455X2 + 2404X4 + 1915X5 - 885.7X11 + 1223X14 + 3284X15 + 2850X18 + 1371X20$$

Lake anchovy (*coilia ectenes*)

$$Y1 = -5146 + 7839X1 + 2240X2 + 1683X4 + 2474X5 - 1036X11 + 1334X14 + 406X15 + 2771X18 + 970.8X20$$

We can tell a fish of these two populations belongs to which one according to above formula. The method was calculating the morphological parameters of a fish, then dividing by its total length to correct, next substituting into above formula to acquire the *Y* value, the bigger one means which population this fish belongs to.

3.3 ANOVA analysis

ANOVA analysis was conducted with all morphological parameters of these two population and showed that there are 7 parameters are extremely significant, they are characters *X*22, *X*16, *X*24, *X*8, *X*6, *X*19, and *X*15; there are 4 parameters are significant, they are characters *X*18, *X*23, *X*10, *X*20. The means and variances of characters of each population and coefficients of difference between two population were showed in tab. 4. We can see from tab. 4 that all coefficients of difference were smaller than the threshold value of subspecies, 1.28, so their difference still limited in different geographical population and have not risen to subspecies level.

Tab. 4 Characters of high variance between two populations of Coilia ectenes

Parameter	Coilia ectenes (Mean ± SD)	Lake anchovy (coilia ectenes) (Mean ± SD)	*F* value	CD (coefficient of difference)
*X*22	0.160 ± 0.017	0.180 ± 0.068	16.0712**	0.233
*X*16	0.527 ± 0.018	0.513 ± 0.050	8.1347**	0.206
*X*24	0.627 ± 0.045	0.619 ± 0.017	7.4092**	0.129
*X*8	0.142 ± 0.015	0.142 ± 0.007	4.7447**	0.000
*X*6	0.050 ± 0.003	0.048 ± 0.006	4.4366**	0.222
*X*19	0.626 ± 0.034	0.604 ± 0.016	4.3335**	0.440
*X*15	0.206 ± 0.006	0.209 ± 0.009	2.3529**	0.200
*X*18	0.683 ± 0.010	0.658 ± 0.015	2.2462*	1.000
*X*23	0.063 ± 0.012	0.064 ± 0.017	2.1217*	0.034
*X*10	0.098 ± 0.011	0.096 ± 0.008	2.0230*	0.105
*X*20	0.035 ± 0.005	0.028 ± 0.004	1.9113*	0.778

** means the differences are extremely significant ($F > F_{39,39,0.01} = 2.297$); * means the differences are significant ($2.297 = F_{39,39,0.01} > F > F_{39,39,0.05} = 1.79$).

4 Discussion

4.1 Application of principal component analysis

Principal component analysis (PCA) is a statistical method to elucidate the differences between populations by synthesizing many parameters into a few factors. Its central idea is to reduce the dimensionality, and it can tell us the parameters which contribute more to their differences of populations according to the value of principle components.

It is an important question to determine how many principal components should be used in principal component analysis. We shall take total variances of sample and eigenvalues into account when determine the number of principle components. Ordinarily, one can conveniently determine the number of principle components by drawing "scree plot of principle components" [6] (Fig. 2). This plot is built according to eigenvalues of principle component and their corresponding serial number.

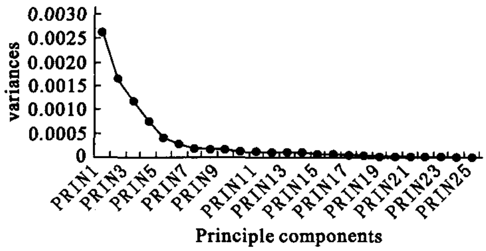

Fig. 2 Scree plot of all components

We can see from the chart (unpublished data) that there is an inflexion at principle component 5 (PC5) in this plot. It means that the eigenvalues of principle components followed PC5 are all smaller and their values are almost the same. So we can choose 4 principle components to synthetize the total variances of all samples effectively.

As an effective multivariate analysis method, pinciple component analysis(PCA) was applied widely in many fields. In fishery field, Wei et al. [7], Yang et al. [8], Ni et al. [9] had studied the morphology of freshwater unionid species, genus Saurogobio, and *Ilisha Elongata*, separately. So we can deeply know the morphological characters of these aquatic biology.

4.2 application of discriminant analysis

Discriminant analysis is a statistical method to be used to look for objective discriminant evidences when we had known the classification[10]. Discriminant analysis has been widely applied to many fields for its strong practicability and can solve actual questions [5]. Stepwise discriminant analysis is a method used mostly. It can build discriminant function by a number of factors which are necessary and best combination from a lot of known factors. Li et al. [10] used this method to predict weather and got good effect. They predicted 12 times of the typhoon which into the predicted areas in 1975, and the accuracy of prediction was 92%.

Discriminant analysis is a common method to be used in identifying fish populations[11]. In fishery field, discriminant function was mainly applied in identifying the mitten crabs from different water systems [12-14] and different strains of tilapia[15,16], and has acquired some significative outcomes. For example, the discriminant formula of mitten crabs from different water systems can help us identify the mitten crabs from different water systems and prevent the intermix of germplasm.

In this study, measureable characters and truss parameters were combined to conduct discriminant analysis of two populations of *Coilia ectenes*, and the effect of discriminant was extremely significant ($P < 0.01$) and the accuracy of discriminant was 100%; The effect of discriminant based on 9 selected morphological characters was also extremely significant ($P < 0.01$), and the accuracy of synthetic discriminant was 97.5%. So it showed that discriminant analysis is an effective method in identifying populations, strains, and subspecies which have near relations.

4.3 Differences between migration group and ecesis group

It is a common phenomenon that a kind of fish has many different groups, such as migration group and ecesis group, with the change of the environments. The former study on the differences between migration group and ecesis group mainly focused on the ecological characters. Take Coilia ectenes as an example, Lake anchovy(*coilia ectenes*) has earlier sex mature age, smaller length of matured fish, minish in matured egg diameter, postpone in the boom period of sprawn, and less plump in liver, than *Coilia ectenes*[1].

It is difficult to identify the migration group and the ecesis group from morphology for they are similar in morphology and their meristic and measureable characters are overlapped. But the quantitative study on their morphological differences have not been reported, this paper is the first report. It revealed from this paper that, there are many difference between migration group and ecesis group, but this difference can not be identified by few index like identifying species, they can be synthesizing discriminated by many parameters.

4.4 the formation of new species

Mayr [17]had some opinion on speciation, "A new species develops if a population which has become geographically isolated from its parental species acquires during this period of isolation characters which promote or guarantee reproductive isolation when the external barries break down". So isolation has important effect on the development of new species. Two population that have the same gene pool at the beginning can acquire difference for different selective pressure in different region inasmuch isolation, also isolation can prevent gene exchange between these two population, so the differences are enlarged till formation of a new species.

But we shall know that the formation of new species by isolation needs long time. *Coilia ectenes* and Lake anchovy (*coilia ectenes*) were the same species originally, but they acquired some differences with the change of environment for isolation. They have some significant changes when the differences accumulated and enlarged. But in morphology, their differences still within different geographical populations and have not risen to subspecies level.

Acknowledgements We thank Mr. Luo Minbo, and Mr. Liu Qizhe for providing help in collecting samples.

References

1 East China Sea Fisheries Research Institute, Chinese Academy of Fisheries Science and Shanghai Fisheries Research Institute. The Fishes of Shanghai Area. Shanghai: Shanghai Scientifical & Technical Publishers, 1990: 93 - 115

2 Mayr E, Linsley E G, Usinger R L. Methods and principles of systematic zoology. New York: McGraw Hill Book Company, 1953: 23 - 39, 123 - 154

3 Brzeski V J, Doyle R W. A morphometric criterion for sex discrimination in tilapia. In: Pullin R S V, Bhukaswan, T Tongutha K *et al* eds. The second inter national symposium on tiapia in aqaculture. ICLARM Conference Proceeding. Bangkok, Thailand. 1988: 439 - 444.

4 Nancy L S, Doyle R W. The coordination of growth in juvenile tilapia(*Oreochromi mossa mbicus* X O. HORNORUM). Proceedings of Second Asian Fisheries Forum, 1989

5 Zhang Y T, Fang K. T. Introduction of multiple analysis. Beijing: Science Press, 1982: 393 - 401

6 Richard A J, Dean W W. Multiple statistical analysis. Beijing: Tsinghua University Press, 2001: 347 - 469

7 Wei K J, Xiong B X. Zhao X H, *et al* Morphological variations and discriminant analysis of five freshwater unionid species (Bivalvia: Unionidae). *Journal of Fisheries of China*, 2003, **27**(1): 13 - 18

8 Yang X P, Zhang M Y, Liu H Z. Studies on morphometrics of the genus Saurogobio. *Acta Hydrobiologica Sinica*, 2003, **27**(2): 164 - 169

9 Ni H E, Chen X. Analysis of shape index system and discriminant of male and female of *Ilisha Elongata*. *Journal of Biomathematics*, 2003, **18**(2): 224 - 228.

10 Ding S C, Multivariate analysis and its appication. Changchun: Jilin People's Publishers, 1981: 259 - 315, 362 - 445

11 Li S F, Wang Q, Wu L Z, *et al* Comprehensive genetic study on Chinese carps. Shanghai: Shanghai Scientific and Technical Publishers, 1990: 145 - 162

12 Xu J W, Ren M R, Li S F. Morphological identification of population of *Eriocheir Sinensis* from changjiang, liaohe, oujiang rivers. *Journal of Fisheries of China*, 1997, **21**(3): 269 - 274

13 Li C H, Li S F. Phylogenesis of populations of mitten crabs (*Eriocheir Sinensis*, *Eriocheir Japonicus*) in six river systems of mainland China : morphology discriminant analysis. *Journal of fisheries of China*, 1999, **23**(4):337 - 342

14 Li Y , Li S F, Wang C H, *et al* Estabilishment and application of morphological discrimination model for juveniles *Eriocheir Sinensis* from Liaohe, Yangtze and Oujiang rivers. *Journal of Fisheries of China*, 2001, **25**(2): 120 - 126

15 Li S F , Li C H, Li J L. Analysis of morphological variations among strains of nile tilapia (*Oreochromis Niloticus*). *Acta Zoologica Sinica*, 1998, **44** (4) : 450 - 457

16 Li J L, Li S F Li Y. Morphology and discrimination of hybrid *Oreochromis Niloticus*(♀) × *Oreochromis Aureus*(♂) and their parents. *Journal of Fisheries of China*, 1999: **23**(3): 261 - 265

17 Mayr E. Systematics and the origin of species. New York: Columbia University Press, 1942: 154 - 173

鲚属两种群的形态变异及综合判别*

程起群[1,2] 韩金娣[1]

(1:中国水产科学研究院东海水产研究所,农业部海洋与河口渔业重点实验室,上海 200090;
2:复旦大学遗传学研究所,上海 200433)

提 要

测量了代表刀鲚种群和湖鲚种群形态特征的传统可量性状和框架参数，采用主成分分析法和逐步判别法，对它们进行形态综合分析．结果表明：两种群的形态差异主要是受额部有鳞部最前缘到背鳍起点的距离所影响的；所有25个形态参数数据判别分析，表明两种群形态差异极显著($P<0.01$)，判别准确率100%；利用挑选后的9项参数判别分析，表明两种群的形态差异极显著($P<0.01$)，判别准确率P1是95% - 100%，P2是95.2% - 100%，综合判别率为97.5%．对各形态参数进行单因子方差分析，计算各参数的差异系数，根据Mayr提出的75%规则，认为它们的形态差异仍然是种内不同地理种群的差异，还没有上升到亚种水平．

关键词 刀鲚(*coilia ectenes*) 湖鲚 (*coilia ectenes*) 判别分析 主成分分析 框架分析 多元分析

分类号 Q7