INTERNATIONAL JOURNAL OF PLANT, ANIMAL AND ENVIRONMENTAL SCIENCES

IJPAES

Volume-3, Issue-1, Jan-Mar-2013 ISSN 2231-4490

Coden : IJPAES www.ijpaes.com

Received: 12Th Jan-2013 *Revised: 18Th Jan-2013* *Accepted: 19Th Jan -2013*

Research Article

PRINCIPAL COMPONENTS ANALYSIS OF *CHONDROCHILUS REGINUS* MORPHOLOGY

Mehdi Yousefian, Yeganeh Yagoubi Kafshkari and Kelara Keshavarz,

Department of Fishery, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran

Corresponding Author: Email: yousefianeco@yahoo.com

ABSTRACT: Tigris, are widespread in west of Iran and lives mostly in river and lake. They are omnivorous fish. This study aims to investigate population structure of *C. reginus* using Principal component Analysis. The study was conducted in three rivers that is the most remote headstream of the Gamasiab River in west part of Iran. Seven quantitative traits were measured for each specimen. After logarithmic transformation of the variables, the PCA first and second axis accounted for 38.4% and 17.6% of the total variance respectively. Considering correlation and contributions, the first axis showed an opposition between morphological variations (length parameters). The second axis showed an opposition between gonad weight and GSI. The third important factor in PCA analysis was the liver weight, HIS and visceral weight that was accounted for 12.6 % of the total variance. In conclusion there were precise differentiations among the three rivers samples in three factor estimated by PCA.
Ket words: *Chondrochilus reginus*, PCA, Gamasib River, HIS, GSI

INTRODUCTION

Natural water such as lake and river are very importanta due to its rule in agricultural and industrial uses. In study of the water resources, investigation of aquatic organism especially fishes are very important. Natural waters have complex interaction of biological, physical and chemical parameters which in suitable environment and condition they provide growth of fish and stable in population. The most important factors influence in existence of the fish in lake or rivers are the presence of enough food and suitable environmental condition that reduce the variation and increase the stable population structure. However stability in natural water is very hard to maintain due to many factors changing the condition of the water. To study the biology and status of a native fish first it is necessary to have the basic information including information on population structure. This information is necessary for the development of management strategies and conserving biodiversity of fish. In study of fish population structure, morphological characters such as morphometric and meristic and molecular analysis have been commonly used to identify stocks of fish [19, 22].

For fish stock management and success in aquaculture expansion and development, studying the native fish and also the habitat they are living, is important. To study the habitat of fish in water ecosystem result in protection and enhance of its stocks. In this respect all the species whether it is economical valuable or not have their important rule in ecosystem.

Up to date not any published data is reported about the population structure of Tigris, ”Orontes ”*Chondrochilus reginus* in east of Iran. Tigris, *Chondrochilus reginus* are widespread in Iraq, Turkey and Iran and lives mostly in river and lake. They are omnivorous fish. In their stomach, we observed worms, insects, eggs and small fishes, algal and vegetables. In Iran they are in separate zone and geography states. They reproduce and adapted in their habitat. Therefore they establish different isolated groups or subpopulation. As noted previously by [8], reproductive isolation among population allows for local adaptation to specific spawning and rearing habitat conditions and results in genetic divergence and wide range of life history characteristics at many geographic scales. These new population may they have in phonotypical and also may they were differed in age at maturity, adult body size and shape, and spawning timing, that we are purchase at present study often vary among populations.

For conservation and aquaculture purpose, analysis of fine scale genetic structure in continuous populations of species is limited by the availability of sufficient markers, therefore, this study aims to investigate population structure of *C. reginus* using Principal component Analysis for morphology characters throughout its distributional range in the Gamasiab Rivers of Iran.
Ecological applications of multivariate statistics have expanded tremendously during the last three Decades. Among these methods, the principal component analysis (PCA) is now used routinely by ecologists [4]. Fish population is used as indicators of ecological integrity. Fish have been shown to be good indicators of larger scale land use and eco-region features [21]. Many researchers have used principal component analysis to describe reach level fish distribution [1]. Inter and intra year trending of environmental variables has also been done using principle component analysis (PCA) ordination [3]. It is known as able to simplify large data sets with reasonable loss of information and to assess inter-correlation among variables of interest [9]. Actually when dealing with high-dimensional data, it is natural to preprocess the data with dimension reduction and feature selection techniques. Principal component analysis (PCA) is a technique of dimension reduction. The importance of a principal component (PC) is proportional to the corresponding eigenvalue, which is the variance of data projected onto this component [11].

MATERIALS AND METHODS

The study was conducted in three river of Dehno, Galae-Gobad and Khachang Rivers that is the most remote headstream of the Gamasib River in west part of Iran. The Gamasiab River is the longest river in Iran originated from a spring with the same name in Nahavand city in West part of Iran. Gamasiab river after passing several province enter the Arvand River and ultimately to Persian Gulf. Geographic attitude of Gamasiab River is *N 34°20´_58.88´ E 47° 59.355´ and* Geographic attitude of Dehno River is *N 34°25.76´ E 48°21.63´*.

Sample Collection

During one year from the summers of 2010, we took sample with electrofished (200 V), in 100 m along the river in closed system (First the upper part and down part) of the sampling site were closed by net and all the fish were collected by 3 times trying using Electro-shocker.
Samples were fixed in formaldehyde at 10% and stored in 70% ethyl-alcohol. Body measurements were taken head to tail using a caliper precision 0.05mm. Seven quantitative traits were measured for each specimen. The following morphological measurements of the distances of landmark in the comparative analysis of Tigris, *C. reginus* populations were used: Fork length, total length, fish weight, standard length, gonado-somatic index (GSI), gonad weight, kidney weight, HIS and visceral weight.
Godado Somatic Index is determined based on following formula:

$$GSI = \frac{W(g)}{W} \times 100$$

Where: W (g) = weight of gonad, W= weight of fish
To determine the age of fish, 5 scales of fish was removed and washed in 4 percent NaCl and checked under binuclear microscope.

Data analyses

Principal components analysis, PCA is used to reduce the dimensionality of data, and to transform interdependent variables into significant and independent components. This statistical method has been extensively described by Gauch, 1982 and Legendre and Legendre, 1998. In the present work, fish species abundances were first log (x+1) transformed in order to satisfy the assumptions of PCA and then submitted to centered and normalized PCA, which reduces the influence of species variation (Doledec and Chessel, 1991). The analysis was therefore conducted on the correlation matrix. PCA was carried out using version 18 of SPSS. For each fish species, individuals recorded young of the year (0+) and older fish one to six year old.

RESULTS

The statistical data of 250 fish sampled in three river at 7 stations are presented in table 1. Number of fish sampled in one trying wearied from 8 to 51 fishes. In total 175 fish were male and 75 fish were female. The fish age were +0 to 6 years old. Therefore there was a large difference between the statistical data including length, gonad weight, GSI and etc.

The PCA allowed the fish species abundances in 7 stations to be taken into account simultaneously aiming to visualize the difference in population structure within the studied area. After logarithmic transformation of the variables, the PCA first and second axis accounted for 38.4% and 17.6% of the total variance respectively (Table 2). A principal component analysis (PCA) performed in isometry free shape space allowed us to explore variation in morphology of *C. reginus*. The results suggested the first two principal components (PC) which explained 56% of the total variance in the data. Considering correlation and contributions, the first axis showed an opposition between morphological variations (length parameters). The second axis showed an opposition between gonad weight and GSI. The third important factor in PCA analysis was the liver weight, HIS and visceral weight that was accounted for 12.6 % of the total variance.
The Mean and Standard deviation of three components are presented in table 3. All coefficients of the first eigenvector were positive, and precisely clustered suggesting it may be interpreted as a multivariate measurement of factors (Table 4). There were precise differentiations among the station samples in three factor estimated by PCA. Patterns of total fish abundance indicated that cyprinids were observed mostly in the Lower Gamasiab River. Total abundance of *C. reginus* was also high in winter. Analysis of spatial trends in the average length and weight of *C. reginus* captured during electrofishing surveys indicated that the average size of this fish was correlated with the weight of fish. Analysis variance of the first component of PCA also indicates significant differences between the stations but for the first component we cannot find any clustering between the station or time of sampling (Table 5). Interesting results are obtain in data of component 2 that the differences in station was significant ($P<0.05$; Table 6). The analysis of variance of component 3 did not showed any significant differences between stations ($P>0.05$).

Table 1- statistical data of sample fish

Parameter	Fork Length (mm)	Standard Length (mm)	Fish weight (g)	Gonad weight (g)	GSI	HSI	Visceral weight (g)
Number of Sample	250	250	250	250	250	231	250
Mean	114.4	107.6	32.3	0.9	1.9	1.1	6.5
Median	124.2	113.2	27.7	0.7	1.4	0.97	6.5
Mode	96.7	40.5 [a]	21.0[a]	0.3	1.00	0.97	1.61[a]
Standard Deviation	46.3	83.3	30.3	0.8	2.0	0.68	3.04
Variance	2140.5	6937.8	920.2	0.7	4.2	0.47	9.3
minimum	11.9	5.1	8.00	0.01	0.04	0.00	-1.5
Maximum	216.5	1222.7	170.8	8.9	21.3	3.4	14.8

Table 2.The percent of alteration by Principle component analysis

Component	Initial Eigenvalues			Extraction sums of squared loading
	Total	% of Variance	Cumulative %	Total
1	3.46	38.40	38.40	3.46
2	1.58	17.60	56.0	1.58
3	1.13	12.00	68.6	1.13

Table 3. The Mean and Standard deviation of component three

Station and time		*First eigen factor*		*Second*		*Third*	
Dehno-Azar89	47	103.8	42.0	0.9	0.7	2.8	0.96
Dehno-Esfand90	46	97.6	25.3	1.1	0.6	2.8	1.1
Ghaleg-Esfand90	51	74.5	42.7	1.3	0.9	2.9	1.2
Ghaleg-Mehr90	13	95.6	26.7	2.5	3.9	2.2	0.8
Ghaleg-Khordad90	10	113.0	26.3	2.4	1.06	2.6	0.6
Dehno-Mehr90	30	87.0	33.3	1.5	1.4	2.8	1.15
Dehno-Mehr89	8	116.0	14.3	0.3	0.4	2.4	1.06
Kharchang-Rriver	45	30.7	3.9	1.9	1.02	2.7	1
Total	250	81.9	41.4	1.4	1.3	2.8	1.09

Table 4. The morphology data were classified and then correlated by PCA

Factors	1	2	3
Fork length	0.964		
Total length	0.937		
Fish weight	0.920		
Standard length	0.635		
GSI		0.910	
Gonad weight		0.853	
Kidney weight			0.798
HIS			0.569
Visceral weight			0.473

Table 5. Duncan multiple range test of first component of PCA

Station	1	2	3	4
Dehno-Mehr89	116.1			
Ghaleg-Mehr90	112.9			
Dehno-Azar89		103.8		
Dehno-Esfand90		97.7		
Ghaleg-Khordad90		95.6		
Dehno-Mehr90			87.1	
Ghaleg-Esfand90			74.5	
Kharchang-Rriver				30.7

Table 6. Duncan multiple range test of second component of PCA

Station	1	2	3
Ghaleg-Khordad90	2.47		
Ghaleg-Mehr90	2.42		
Kharchang-Rriver		1.86	
Ghaleg-Esfand90		1.56	
Dehno-Mehr90		1.28	
Dehno-Esfand90		1.14	
Dehno-Azar89			0.938
Dehno-Mehr89			0.338

DISCUSSION

The study of population structure and variation among the fish is a branch of biology that also IUCN believe it should be consider and the genetic variation should remain and protect. This is because the genetic variation is the basic for future animal evolution and adaptation is the way for animal inheritance and heredity. Therefore the protection of biotic diversity and ecosystem safety is very important.

The result of present study showed the fish of different sample station are significantly different in several characteristic, especially in meristic parameters and GSI as well. For analyzing of the data we used PCA method. Previously Genetic differentiation within and between populations was analyzed using principal component analysis (PCA) on the SNP genotype data [14]. This approach detects and summarizes differentiation by identifying a set of successive orthogonal principal components (PCs), each explaining a higher amount of the total variation than any remaining PCs. Significance is determined by comparing eigenvalues corresponding to successive PCs to the theoretical Tracy–Widom (TW) distribution of largest eigenvalues [18]. In the present study the initial eigenvalues was scored 38.4, 17.6 and 12 as percent of variances. It consists of more than 68 % of total variance. The main reason for using PCA in present study was that in scoring the morphological data numerous data was collected from high number of fish (250 fishes) and it make difficulties in analysis of data.

Handling the large number of data and analyzing of different characteristics is complicated. In PCA, unusual observations are frequently deleted from the data sets to reduce data heterogeneity [5]. PCA are used in a series of ecological studies e.g. [20] used this method to study the population of fish. In their study diversity of the migratory environmental history of freshwater resident flathead mullet *Mugil cephalus* was examined by analyzing the otolith elemental composition of mullet collected in midstream of the Tanshui River, northern Taiwan by laser-ablation inductively coupled plasma mass spectrometry. Also [15] for finding the distribution pattern of larval fish populations in the Northwestern Mediterranean used PCA method. [11] denoted that handling genotype data typed at hundreds of thousands of loci is very time-consuming and it is no exception for population structure inference.

PCA is a good method for analyzing the distribution of *Chondrostona regium* in west of Iran, because this fish is distributed in several river with environmental condition. Numerous sampling and information make it difficult to analysis the differences. This method also were used by Brosse et al., 2001 that had a study to analyses the spatial occupancy of several European freshwater fish species in the littoral zone of a large French lake. The data matrix consisting of 710 samples and 15 species was analyzed using PCA and SOM. Both methods provided insights on the major trends in fish spatial occupancy.

[10], in their study of Tree population structure, regeneration and expected future composition at different levels of *Lantana camara* L. invasion in the Vindhyan tropical dry deciduous forest of India explained in PCA component are successively extracted from a matrix similarities. In PCA all individuals contribute equally to the component, avoiding dominance of outliers. Mathematically, PCA involves eigen analysis of a symmetric matrix to similarities to produce a series of eigen values and there corresponding eigen vectors [13]. There are as many eigen values as there are rows (or columns) in the matrix and conceptually they can be considered to measure the strength (relative length) of an axis. Each eigen value has an associated eigen vector. An eigen value gives the length of an axis; the eigen vector determines its orientation in space [10]. In present study, the main eigen value consist of morphometric characters showing the differences in length and weight of fish at different rivers. The fish of up to one year was very high in Kharchang River. This seems to be related to the condition of river and availability of food in this river. The Kharchang River is elongated in lower field comparing to the two other rivers which is located in altitude. Therefore the river is riches in grasses and other nutrient and higher temperature which is good for propagation and rearing of fish. This result is accordance with the work of [4]. They found higher fingerling and stated the abundance due to food and shelter that the young fish can escape from the enemy. A key factor facilitating divergence in fish is the stability and location of feeding habitats. If pronounced and persistent ecological differences occur among feeding habitats, subsequent behavioral and morphological specialization to these habitats is more likely [16].

In assigning individual relationship some markers (or variants) are useful and more informative than others In present study component 2 and 3 showed specific pattern that is used illustration of different sampling groups. Selecting and using only the most informative markers for population assignment can reduce both time and genotyping costs while retaining most of the power of the complete set of markers. In sampling *C. regium* in different river, GSI had a component separate from the other two components and it showed that can be used as a special marker for evaluation of population structure or shows the habitats differences.

By analyzing the morphology of fish it was showed that the fish in Kharchang River was frequent and small, it is likely related to increased survival of eggs, fry, and juveniles in the relatively warmer downstream of the River. Higher water temperature has a significant influence on *C. regium* behavior and growth. *Abudefduf saxatilis* populations along South America coast showed a clonal gradient in the north south direction in relation to their meristic variables. Temperature may be a decisive factor explaining regional differences among these samples and it is not uncommon for meristic traits to have lower number of segments in colder waters at the southern coast, when compared to northern warmer waters (Wagner et al., 2006). Usually, parameters delaying ontogenetic development, such as hypoxia, low temperatures, high salinity and reduced food sources, favor the development of greater number of body segments [2, 17].

In conclusion there were precise differentiations in *C. reginus* among the three rivers samples in three factor estimated by PCA, however it is suggested to have molecular analysis of the same fish in different river of the west of Iran to make the study complete.

REFERENCES

[1]. Bain, M.B. and J.G. Knight 1996. Classifying Stream Habitat Using Fish Community Analyses. Pages B106-B117 in Proceedings of the second IAHR Symposium on Habitat Hydraulics, Ecohydraulics 2000. Institute National de la Recherche Scientifique - Eau, Ste-Foy, Quebec, Canada. 1996.

[2]. Barlow, G.W. 1961. Causes and significance of morphological variation in fishes. Systematic Zoology, 10: 105-117

[3]. Bartell S.M., Allen T.F.H. and J.F. Koonce 1978. An assessment if principle components analysis of phytoplankton periodicity in Lake Wingra. Phcologia, 17:1-11

[4]. Brosse S., Giraudel J.L. and S. Lek. 2001. Utilization of non-supervised neural networks and principal component analysis to study fish assemblages. Ecological Modelling 146: 159–166.

[5]. Copp G.H., Guti G., Rovny B. and J. Cerny 1994. Hierarchical analysis of habitat use by 0+ juvenile fish in Hungarian/Slovak flood plain of the Danube river. Env. Biol. Fish, 40:329–348.

[6]. Doledec S. and D. Chessel 1991. Recent developments in linear ordination methods for environmental sciences. In: Council of Science Research Integration (Ed.), Trends in ecology. Research trends Publishers, India. pp. 1–21.

[7]. Gauch H.G. 1982. Multivariate Analysis in Community Ecology. Cambridge University Press, Cambridge.

[8]. Glauflin M.T. 2010. Influences of spawning habitat and geography: population structure and juvenile migration timing of sockeye salmon (*Oncorhynchus nerka*) in the Wood River lakes, Alaska. A thesis of M.Sc. University of Washington. P,50

[9]. Grossman G.D., Nickerson D.M. and M.C. Freeman 1991. Principal component analyses of assemblage structure data: utility of tests based on eigenvalues. Ecology, 72: 341–347.

[10]. Gyan P.S. and A. S. Raghubanshi 2006. Tree population structure, regeneration and expected future composition at different levels of *Lantana camara* L. invasion in the Vindhyan tropical dry deciduous forest of India. Lyonia,11(1): 27-39.

[11]. Lee C., Abdool A. and C.H. Huang 2009. PCA-based population structure inference with generic clustering algorithms. doi: 10.1186/1471-2105-10-S1-S73

[12]. Legendre, P., Legendre, L., 1998. Numerical Ecology. Elsevier. Amsterdam, 853p.

[13]. Marshall S. and M. Elliot 1998: Environmental influences on the fish assemblage of the Humber Estuary. U.K. Estuarine, Coastal and Shelf Science 46: 175–184.

[14]. Patterson N.J., Price A.L. and D. Reich 2006. Population structure and eigenanalysis. PLoS Genet. 2:190.

[15]. Sabates A. 1990. Distribution pattern of larval fish populatlons in the Northwestern Mediterranean. Mar Ecol Prog Ser., 59:85-82.

[16]. Skúlason S., Snorrason S.S. and B. Jonsson 1999. Sympatric morphs, populations and speciation in freshwater fish with emphasis on *arctic charr*. In: Magurran AE, May RM, eds. Evolution of biological diversity. New York, NY: Oxford University Press, 70–92.

[17]. Shaklee, J. B. C.S. Tamaru, C. S. 1981. Biochemical and morphological evolution of Hawaiian bonefishes (Albula). Systematic Zoologic, 30: 125-146.

[18]. Tracy C.A. and H. Widom 1994. Level-spacing distributions and the airy kernel. Communications in Mathematical Physics, 159: 151–174.

[19]. Turan C., Erguden D., Turan F. and M.Gurlek 2004: Genetic and Morphologic Structure of *Liza abu* (Heckel, 1843) Populations from the Rivers Orontes, Euphrates and Tigris. Turk J. Vet. Anim. Sci. 28:729–734.

[20]. Wagner F.M., Shibatta O.A. and P.M. Ga;etti-Jr 2006. Multivariate morphological analyses in continental and island populations of *Abudefduf saxatilis* (Linnaeus) (*Pomacentridae, Perciformes*) of Western Atlantic Pan-American Journal of Aquatic Sciences, 1 (2): 49-56

[21. Whittier T.R., Hughes R.M. and D.P. Larsen 1988. Correspondence between ecoregions and spatial patterns in stream ecosystems in Oregon. Can. J. Fish. Aquat. Sci. 45:1264-1278.

[22]. Yousefian, M. Genetic Variations of Common Carp (*Cyprinus carpio* L.) In South-Eastern Part of Caspian Sea Using Five Microsatellite Loci World Journal of Zoology, 6 (1): 56-60, 2011