INNSPUB

Journal of Biodiversity and Environmental Sciences (JBES)
ISSN: 2220-6663 (Print), 2222-3045 (Online)
http://www.innspub.net
Vol. 6, No. 3, p. 200-206, 2015

RESEARCH PAPER | OPEN ACCESS

Agglomerative hierarchical clustering based on morphmetric parameters of the populations of *Labeo rohita*

Fayyaz Rasool[1*], Shakeela Parveen[1], Ehsan Mahmood Bhatti[2], Noor Khan[1]

[1]*Faculty of Fisheries and Wildlife, Department of Fisheries and Aquaculture, University of Veterinary and Animal Sciences, Lahore, Pakistan*

[2]*Fisheries Research and Training Institute, Manawan-Lahore, Department of Fisheries Punjab, Pakistan*

Key words: AHC, *Labeo rohita*, Hatchery, Riverine, Morphometric.

Article published on March 17, 2015

Abstract

Labeo rohita populations from five geographical locations from the hatchery and Riverine system of Punjab-Pakistan were studied for the clustering on the basis of similarities and differences based on morphometric parameters within the species. Agglomerative Hierarchical Clustering (AHC) was done by using Pearson Correlation Coefficient and Unweighted Pair Group Method with Arithmetic Mean (UPGMA) as Agglomeration method by XLSTAT 2012 version 1.02. A dendrogram with the data on the morphometrics of the representative samples of each site divided the populations of *Labeo rohita* in to five major clusters or classes. The variance decomposition for the optimal classification values remained as, 19.24% for within class variation while 80.76% for the between class differences. The representative central objects of the each class, the distances between the class centroids and also distance between the central objects of the classes were generated by the analysis. A measurable distinction between the classes of the populations of the *Labeo rohita* was indicated in this study which determined the impacts of changing environment and other possible factors influencing the variation level among the populations of the same species.

***Corresponding Author:** Fayyaz Rasool ✉ fayyazrasool@uvas.edu.pk

Introduction

Labeo rohita is popular in Pakistan, India, Bangledesh and Thailand. It is a non-oily or white fish. It is commonly known as Rohu or Dumbra. It is a fish of the carp family, *Cyprinidae* found commonly in rivers and freshwater lakes in and around South Asia and South-East Asia. This is omnivorous in nature (Keralaagriculture, 2012). The studies have been conducted for the comparison of morphometric parameters along with genetic disturbances for the conservation of this endemic species in the sub-continent (Faith *et al.* 2004). (*Labeo rohita,* commonly known as rohu is a very good model in aquaculture due to faster growth rate. Indian major carps account for approximately 75% of the total inland aquaculture production in India (Barman *et al.,* 2003).

El-Zaeem *et al.,* (2012), conducted a study on the variation in phenotype based on morphometric character indices and meristic counts among different wild and cultured Nile tilapia (*Oreochromis niloticus*). The results revealed that there were significant differences ($P \leq 0.05$) in most of morphometric character indices and meristic counts among different wild and cultured Nile tilapia populations tested. The hierarchical cluster analysis based on each phenotype and genotype analysis grouped the four populations into two major category groups: Edku Lake, Manzalah Lake and Nile river populations group and cultured population group. Within these major grouping, wild Nile tilapia were grouped close together. Also, Edku Lake population appears to be more similar to that of Manzalah Lake population than that of Nile river population.

The clustering analysis by Unweighted Pair Group Method with Arithmetic Means, for the genetic diversity amongst the natural and hatchery raised populations of Indian major carp *Calta catla* were studied. The 30-samples for each natural population were collected from the Halda, Jamuna and Padma Rivers and also from one hatchery population in Bangladesh. The geographic distances were directly correlated with the distances measured in gene diversity appeared in the dendrogram. dendrogram showed 2-clusters, the population of River Halda appeared in one cluster other all natural populations including hatchery bunched in the second cluster. This indicates that first cluster from Halda was qualitatively different from the remaining populations. (Rahman *et al.,* 2009). Rana *et al.,* (2004) conducted a study with two hatchery and one river populations of *Labeo rohita* and *Catla catla* from different geographical locations in India. From their results they concluded that the *C. catla* population from Yamuna River Delhi, which is wild population, is highly diverse and is not included in any cluster. The Powerkheda population of *C. catla* placed in the same cluster consisting of *L. rohita* population as in the UPGMA tree. They postulated that mean diversity at population level is nearly 10% in *C. catla* and 1.33% at *L. rohita* level with 24 and 62%. Silva (2003) conducted a study on the morphometric variation among sardine (*Sardina pilchardus*) populations from the northeastern Atlantic and the western Mediterranean. Their results indicated that there was also some evidence that fish from the western Mediterranean and the Azores form a separate morphometric group. These results question both the homogeneity within the Atlanto-Iberian sardine stock and the validity of its current boundaries.

Chauhan *et al.* (2007) to investigate the different populations of *Cirrhinus mrigala* from different Riverine sites in India. *Cirrhinus mrigala* were collected from ten Riverine sites from the Indus, Ganges, Brahmaputra and Mahanadi basins. AMOVA analysis results showed that there is low differentiation among sub-populations. From these results it can be concluded that this low level of differentiation is may be due to the common ancestors of the populations in the pre-historic time and possible exchange of representatives between wild populations of *C. mrigala* in different River basins.Results of studies on *C. carpio* by Dayu *et al.,* (2007) also indicated that there was a correlation between the clustering result and the geographical distribution. In Pakistan the such work, which based on the morphometeric parameters for differentiation among the different varieties of plants of the same species has been conducted (Nisar *et al.,* 2010), mahaseer fish populations (Pervaiz *et al.,* 2012) and populations of *Heterodera zeae,* chickpea and walnut varieties (Abdollahi, 2009; Talebi *et al.,* 2008 and Asadian, 2005) in Iran, but the work on fish or other aquatic animals in small counts. Therefore, this study was planned to sort out variations among the populations of *Labeo rohita* from the same species but from the different geographical locations.

Materials and methods

Samples and Sampling Sites

The hundred *Labeo rohita* from each site were collected i.e. were the representatives of hatchery raised and natural populations of the Riverine systems of Punjab viz., UVAS-Fish Hatchery, C-block Ravi campus Pattoki District Kasur, Trimu Barrage at the junction of River Chenab and Jhelum near district Jhang, Taunsa Barrage at River Indus near Tehsil Kot Adu District Muzaffar Garh, Qadirabad Barrage at River Chenab near District Mandi Bahuddin and Baloki Barrage at River Ravi near Tehsil Bhai Phero District Kasur.

Data Recording

Data regarding the morphogenetic parameters viz., body weight, fork length, total length and lengths of dorsal, caudal, anal, pectoral and pelvic fins of each individual were recorded.

Statistical Analysis

Twenty representative samples collected from each geographical location and selected randomly from the total hundred samples of each site were used in clustering analysis. Agglomerative Hierarchical Clustering (AHC) was done by using Pearson Correlation Coefficient and Unweighted Pair Group Method with Arithmetic Mean (UPGMA) as Agglomeration method by XLSTAT 2012 version 1.02.

Table 1. Summary statistics of the Morphometric Parameters.

Variable	Min.	Max.	Mean	Std. deviation
Weight	180.000	345.000	246.850	31.241
F-Length	20.800	26.900	24.108	1.254
T-Length	23.900	28.800	26.953	1.238
D-Fin	4.800	8.100	6.348	0.639
C-Fin	3.700	6.400	5.470	0.589
A-Fin	3.100	5.800	4.414	0.570
Pe-Fin	3.000	5.400	4.477	0.519
Pl-Fin	2.900	5.500	4.253	0.544

Results

Clusters/Classes

A dendrogram showing clustering of similar individuals based on eight most correlated parameters is presented here for the documentation of the results about the populations of *L. rohita*. The dendrogram divided the populations of *L. rohita* in to five major clusters or classes (Fig. 1). The first class/cluster encompasses of 55 individuals of same characteristics. In the second class/cluster 7 individuals of the same characteristics, in the third class/cluster23 individuals with similar morphometrics, in fourth class/cluster14 individuals with same characters and in the fifth class/cluster only one individual were grouped together (Fig. 2 and Table 2). The division of all randomly selected 100 *L. rohita* samples, 20 individuals for each sampling sites was as follows in different five classes. Amongst the 55 individuals of the same characters in the first class/cluster; the 8 individuals were from the Hatchery samples, 7 individuals were from the Indus River samples collected from Taunsa barrage, 11individuals included in this class/cluster were from the junction of Chenab and Jhelum Rivers at Trimu barrage, the 17 individuals were from the Chenab River collected from Qadirabad barrage while 12 individuals in this first class/cluster were from the River Ravi sampled from Baloki barrage. From the 7 individuals of the same characters in the second class/cluster; 1 individual was from the Hatchery samples, 2 individuals were from the Indus River samples collected from Taunsa barrage, 1 individual included in this class/cluster was from the junction of

Chenab and Jhelum Rivers at Trimu barrage, in this class/cluster there was not any single individual from the Chenab River collected from Qadirabad barrage while 3 individuals in this first class/cluster were from the River Ravi sampled from Baloki barrage. The 23 individuals of the same characters in the third class/cluster were comprised of 10individuals from the Hatchery samples, 4 individuals from the Indus River samples collected from Taunsa barrage, 5 individuals from the junction of Chenab and Jhelum Rivers at Trimu barrage, the 3 individuals from the Chenab River collected from Qadirabad barrage while 1individualamongst the samples collected from the River Ravi sampled from Baloki barrage. Amongst the 14 individuals of the same characters in the fourth class/cluster; 1 individual was from the Hatchery samples, 6individuals were from the Indus River samples collected from Taunsa barrage, 3 individuals included in this class/cluster were from the junction of Chenab and Jhelum Rivers at Trimu barrage, there was not individual in this group from the samples which were collected from the Chenab River from Qadirabad barrage while 4individuals in this class/cluster were from the River Ravi sampled from Baloki barrage. In the fifth and last class/cluster there was only one individual having some unique characteristics and it was sampled from Indus River from the Taunsa barrage (Table 2).

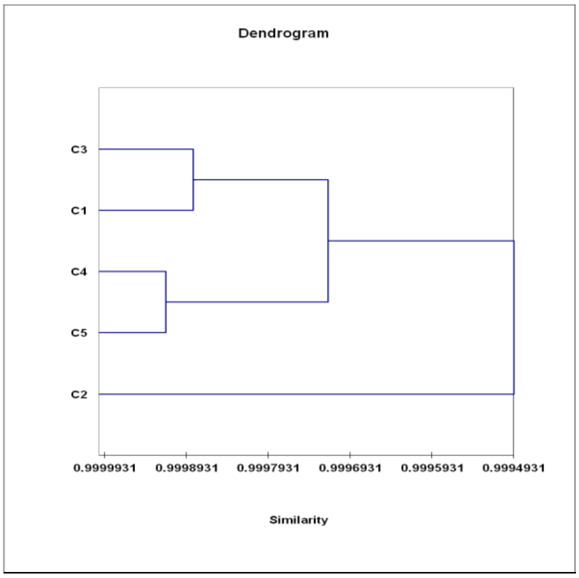
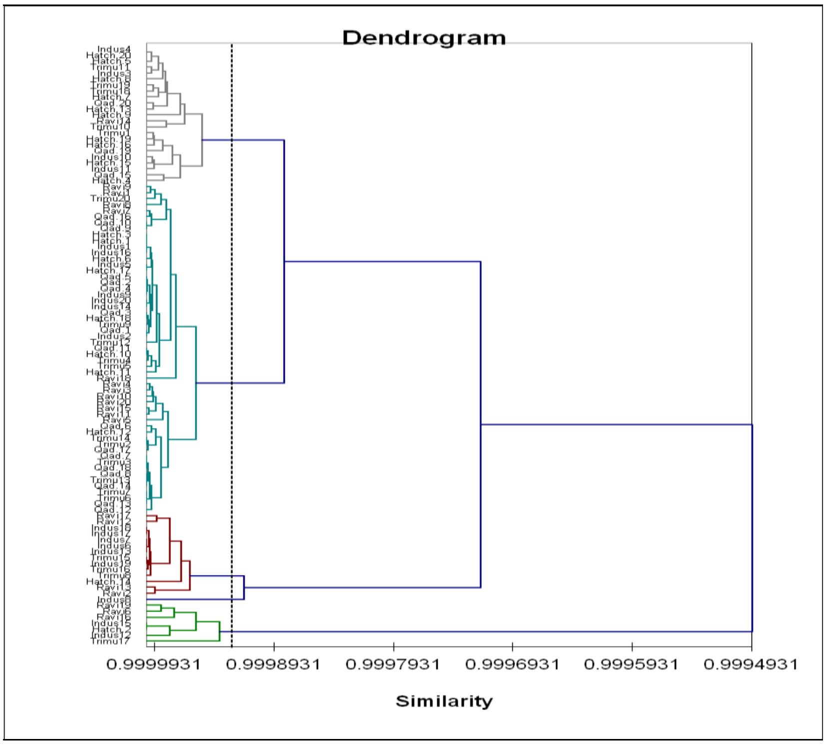


Fig. 1. Dendrogram showing division into classes.



Note: Hatch. (Samples from Hatchery), Trimu (Samples from Trimu Barrage), Indus (Samples from River Indus, Taunsa Barrage), Qad. (Samples from Qadirabad Barrage) and Ravi (Samples from Ravi Barrage)

Fig. 2. Dendrogram showing division of individuals in each class.

Table 2. Results by class.

Class	1	2	3	4	5
Objects	55	7	23	14	1
Sum of weights	55	7	23	14	1
Within-class variance	134.703	442.258	249.041	194.145	0.000
Minimum distance to centroid	1.757	4.113	2.579	2.801	0.000
Average distance to centroid	10.249	18.487	13.404	9.610	0.000
Maximum distance to centroid	26.763	22.403	27.615	35.069	0.000

Individuals of Cluster/Class First

Hatch.1,Hatch.3,Hatch.6,Hatch.10,Hatch.11,Hatch.12,Hatch.17,Hatch.18,Indus1,Indus2,Indus5,Indus9,Indus14,Indus16,Indus20,Trimu2,Trimu3,Trimu4,Trimu5,Trimu6,Trimu7,Trimu9,Trimu12,Trimu13,Trimu14,Trimu20,Qad.1,Qad.2,Qad.3,Qad.4,Qad.5,Qad.6,Qad.7,Qad.8,Qad.9,Qad.10,Qad.11,Qad.12,Qad.13,Qad.14,Qad.16,Qad.17,Qad.18,Ravi1,Ravi3,Ravi4,Ravi5,Ravi7,Ravi8,Ravi9,Ravi10,Ravi11,Ravi15,Ravi18,Ravi20

Individuals of Cluster/Class Second

Hatch.2, Indus12, Indus15, Trimu17, Ravi6, Ravi16, Ravi19

Individuals of Cluster/Class Third

Hatch.4,Hatch.5,Hatch.7,Hatch.8,Hatch.9,Hatch.13,Hatch.15,Hatch.16,Hatch.19,Hatch.20,Indus3,Indus4,Indus10,Indus11,Trimu1,Trimu10,Trimu11,Trimu18,Trimu19,Qad.15,Qad.19,Qad.20,Ravi14

Individuals of Cluster/Class Fourth

Hatch.14, Indus6, Indus7, Indus13, Indus17, Indus18, Indus19, Trimu8, Trimu15, Trimu16, Ravi2, Ravi12, Ravi13, Ravi17

Individuals of Cluster/Class Fifth

Indus8

Variance decomposition for the optimal classification

The variance decomposition for the optimal classification values remained as, 19.24% for within class variation while 80.76% for the between class differences (Table 3). The class centroid for the each characteristics resemblance for clustering was also generated (Table 4). The distance between the class/cluster centroids remained as; 48.678 for class one and two, 32.777 for class one and three, 46.391 for class one and four, 94.687 for class one and five, 15.977 for class two and three, 95.064 for two and four, 143.355 for two and five, 79.161 for three and four, 127.455 for three and five while this distance between class four and five centroids was 48.300 (Table 5). The central objects of the five classes were as; Hatch.10 for class first, Trimu17 for class second, Hatch.5 for class three, Trimu8 for class four and for class five the central object was Indus8 (Table 6). The distances between the central objects of the different cluster/classes remained were also calculated (Table 7).

Table 3. Variance decomposition for the optimal classification.

	Absolute	Percent
Within-class	188.740	19.24%
Between-classes	791.997	80.76%
Total	980.737	100.00%

Table 4. Class centroids for each class.

Class	Weight	F-Length	T-Length	D-Fin	C-Fin	A-Fin	Pe-Fin	Pl-Fin
1	250.345	24.262	27.120	6.433	5.467	4.505	4.524	4.322
2	201.714	24.186	26.214	5.857	4.571	3.429	3.700	3.457
3	217.609	23.204	26.178	6.083	5.461	4.074	4.187	3.878
4	296.714	24.871	27.814	6.636	5.914	5.043	5.093	4.936
5	345.000	25.200	28.700	7.200	5.900	5.300	5.400	5.100

Table 5. Distances between the class centroids.

	1	2	3	4	5
1	0	48.678	32.777	46.391	94.687
2	48.678	0	15.977	95.064	143.355
3	32.777	15.977	0	79.161	127.455
4	46.391	95.064	79.161	0	48.300
5	94.687	143.355	127.455	48.300	0

Table 6. Central objects.

Class	Weight	F-Length	T-Length	D-Fin	C-Fin	A-Fin	Pe-Fin	Pl-Fin
1 (Hatch.10)	251.000	24.600	28.100	7.100	6.100	5.300	4.800	4.500
2 (Trimu17)	198.000	25.000	25.200	6.900	4.900	3.700	3.900	3.800
3 (Hatch.5)	220.000	23.100	26.300	6.800	5.900	4.300	4.500	4.100
4 (Trimu8)	294.000	25.000	27.500	6.900	6.300	5.000	5.200	5.300
5 (Indus8)	345.000	25.200	28.700	7.200	5.900	5.300	5.400	5.100

Table 7. Distances between the central objects.

	1 (Hatch.10)	2 (Trimu17)	3 (Hatch.5)	4 (Trimu8)	5 (Indus8)
1 (Hatch.10)	0	53.131	31.111	43.017	94.008
2 (Trimu17)	53.131	0	22.150	96.067	147.068
3 (Hatch.5)	31.111	22.150	0	74.052	125.053
4 (Trimu8)	43.017	96.067	74.052	0	51.019
5 (Indus8)	94.008	147.068	125.053	51.019	0

Discussion

The dendrogram developed by Agglomerative Hierarchical Clustering (AHC) by using Pearson Correlation Coefficient and Unweighted Pair Group Method with Arithmetic Mean (UPGMA) for populations of *L. rohita,* divided the genotypes in to five major clusters or classes. Each class almost contains the samples from the different geographical samples except the last class where there was only one individual having some distinction. These results are in accordance with the findings of El-Zaeem *et al.,* (2012), who conducted a study on the variation in phenotype based on morphometric character indices and meristic counts among different wild and cultured Nile tilapia (*Oreochromis niloticus*) and found that hierarchical cluster analysis based on each phenotype and genotype analysis grouped the four populations into two major category groups. These results indicated that the one sample from river Indus which is a natural/wild representative was in separate group. These results are confirmation of the results postulated by Rana *et al.,* (2004) who conducted a study with two hatchery and one river populations of *Labeo rohita* and *Catla catla* from different geographical locations in India. From their results they concluded that the *C. catla* population from Yamuna River Delhi, which is wild population, is highly diverse and is not included in any cluster. The variance values remained as, 19.24% for within class variation while 80.76% for the between class differences. These results are in accordance with the results of Chauhan *et al.* (2007) who studied different populations of wild *C. mrigal* from different River basins and concluded that there exist low level of differentiation between the populations of the same species and this may be due to common ancestry and exchange of individuals among the River basins. The results are also confirmation of the results indicated in the study conducted by Dayu *et al.*(2007) on the genetic similarity amongst the wild populations of *Cyprinus carpio* and concluded that there was a correlation between the clustering result and the geographical distribution.

References

Abdollahi M. 2009. Hierarchical cluster analysis of Indian populations of *Heterodera zeae* based on second stage juveniles and egg morphometrics. International Journal of Agriculture and Biology **11**, 756–760

Asadian G, Pieber K. 2005. Morphological Variations in Walnut Varieties o f the Mediterranean Regions. International Journal of Agriculture and Biology **7(1)**, 71-73

Barman HK, Barat A, Bharat M, Yadav M, Banerjee S, Mehar S, Reddy PVGK, Jana RK. 2003. Genetic variation between four species of Indian major carps as revealed by random amplified polymorphic DNA assay. Aquaculture **217,** 115-123.

Chauhan T, Lal KK, Mohindra V, Singh RK, Punia P, Gopalakrishnan A, Sharma PC, Lakra WS. 2007. Evaluating genetic differentiation in wild populationsof the HIndian major carp, Cirrhinus mrigala (Hamilton–Buchanan, 1882): Evidence from allozyme and microsatellite markers. Aquaculture **269,** 135–149.

Dayu LK, Dahai Y, Qianqian S, Xiaowen, Liqun L. 2007. Microsatellite DNA marker analysis of genetic diversity in wild common carp (*Cyprinus carpio* L.) populations. J. of Genet. and Geno., **34 (11),** 984-993.

El-Zaeem SY, Ahmed MMM, Salama ME, Abd-El-Kader WN. 2012. Phylogenetic differentiation of wild and cultured Nile tilapia (*Oreochromis niloticus*) populations based on phenotype and genotype analysis. African Journal of Agricultural Research **7(19),** 2946-2954

Faith DP, Reid CAM, Hunter J. 2004. Integrating phylogenetic diversity, complementarily, and endemism for conservation assessment. Conservation Biology **18,** 255-261.

Kerelaagriculture. 2012. "Composite fish culture". http://www.Kerelaagriculture.gov.in.

Nisar MF, Hussain K, Majeed A. 2010. Analysis of endosperm proteins and genetic diversity in Pakistani wheat varieties. International Journal of Agriculture and Biology **12(2),** 261–265

Pervaiz K, Iqbal Z, Mirza MR, Javed MN, Naeem M. 2012. Meristic and morphometric studies on indus mahseer *Tor macrolepis*(Teleostei: Cyprinidae) from district Attock, Pakistan. International Journal of Agriculture and Biology **14,** 169–175

Rahman SMZ, Mukhlesur RK, Shahidul I, Samsul A. 2009. Genetic variation of wild and hatchery populations of the catla Indian major carp (*Catla catla* Hamilton 1822: *Cypriniformes, Cyprinidae*) revealed by RAPD markers. Genetics and Molecular Biology **32,** 197-201.

Rana RS, Bhat KV, Lakhanpal1 S, Lakra WS. 2004. Comparative Genetic Diversity in Natural and Hatchery Populations of Indian Major Carps (*C. catla* and *L. rohita*). Asian-Aust. Journal of Animal Science **17(9),** 1197-1203

Silva A. 2003. Morphometric variation among sardine (*Sardina pilchardus*) populations from the northeastern Atlantic and the western Mediterranean. ICES Journal of Marine Science **60,** 1352-1360

Talebi R, Fayaz F, Mardi M, Pirsyedi SM, Naji AM. 2008. Genetic relationships among chickpea (*Cicer arietinum*) elite lines Based on RAPD and Agronomic Markers. International Journal of Agriculture and Biology **10,** 301-305