Similarity and genetic relationship analysis of 28 species of Pangasiidae (Siluriformes, Ostariophysi)

R Gustiano¹, V A Prakoso¹,²*, M H F Ath-thar¹, I I Kusmini¹, and D Radona¹
¹Research Institute for Freshwater Aquaculture and Fisheries Extension, Ministry of Marine Affairs and Fisheries, 16129 Bogor, West Java, Indonesia
²Corresponding author: vitas.atmadi@gmail.com

Abstract. Pangasiid catfishes have most diverse morphotype among members in the family. In the past, misidentification often occurred for the species definition. After numerical taxonomy developed, the level of similarity or phenetic analysis was used to distinguish different taxa. While relationship and evolutionary evidence within species in Pangasiidae was analyzed based on the genetic data. The aim of the present study was to understand the congruency between phenetic and genetic analysis in family Pangasiidae. Euclidean distance was applied on phenetic analysis to produce phenogram of similarity. Sequencing of 12s rDNA was conducted to illustrate the genetic relationship. Both analyses showed that species member in Pangasiidae supported the genera proposed in the previous study.

1. Introduction
Pangasiids catfishes exist in the main freshwater river from South Asia to Southeast Asia region. This group of catfishes is very important for fisheries for a long time in many Asian countries [1]. At present time, three species, Pangasius hypophthalmus, P. boucorti, and P. djambal have been utilized widely in aquaculture production in Southeast Asia [2]. Even though, for another promising species were also observed such as P. kunyit and P. nasutus [3], the breeding program is not established yet. After Vidthayanon [4], a comprehensive study found seven new species described [5,6]. They are P. rheophilus [7], P. mahakamensis and P. mekongensis [8], P. kunyit and P. sabahensis [9], and Helicophagus leptorhyncus [10]. Morphologically, the genera of pangasiid catfishes recognized from each other based on number of pelvic fin rays, teeth formation, snout and nostril position, length of barbels, and eye diameter [11]. Inside the genera, members are differentiated according to the existing variation [10,12]. For this purpose, the phenetic analysis was studied to elaborate the difficulty identification of different species in field. Partial phylogenetic of Pangasiidae has been studied by Pouyaud et al. [13] using allozyme, cyt-b and D-loop. Further established works were done by Gustiano [1], Pouyaud et al. [14], Gustiano & Pouyaud [5], Karinthanyakit and Jondeung [15], and Azlina et al. [16]. Most of the studies confirmed the recognition the validity of sub-genera and, or species groups existed. Moreover, there are a congruency among the phylogenetic trees resulted from trees from the osteological [1], molecular [14], and biometric analyses [17]. Above studies also confirm geological history, paleontology hypotheses, and proposed models of the different taxa in the observed areas [1]. The genus Pangasius is monophyletic, in which may already existed before separating islands with Asian continent 20 million years ago [14].
Since technology of DNA sequencing developed, molecular information is very useful to clarify the phylogenetic among taxa exist and their history. However, relationship among higher taxa still needs appropriate part of DNA for sequencing. The present study was to understand the congruency between phenetic and genetic analysis using 12SrDNA among 28 species in family Pangasiidae.

2. Material and methods
The 999 specimens of 28 species pangasiids catfishes were used for the conducted study. Measurement taken on each specimen, 35 characters, followed the illustration given by Gustiano [18] (Figure 1).

![Figure 1](image.png)

**Figure 1.** Measurements taken on *Pangasius* specimens: 1. Standard length; 2. Head length; 3. Snout length; 3a. Anterior snout width; 3b. Posterior snout width; 4. Head depth; 5. Head width; 6. Predorsal length; 7. Caudal peduncle length; 8. Caudal peduncle depth; 9. Pectoral fin length; 10. Pectoral spine length; 11. Dorsal fin length; 12. Dorsal spine width; 13. Pelvic fin length; 14. Anal fin height; 15. Anal fin length; 16. Adipose fin height; 17. Adipose fin width; 18. Eye diameter; 19. Mouth width; 20. Lower jaw length; 21. Interorbital length; 22. Distance snout to isthmush; 23. Postocular length; 24. Maxillary barbel length; 25. Mandibulary barbel length; 26. Body width; 27. Prepectoral length; 28. Prepelvic length; 29. Vomerine length; 30. Vomerine width; 31. Palatine length; 32. Palatine width; 33. Dorsal spine width.

The mean values of each measured characters were used for the phenotypic analysis. The data coded on a scale of 0-4, with 0=x < (mean-sd); 1=[(mean-sd) ≤ x < mean]; 2=[mean ≤ x < (mean+sd)]; 3=[(mean+sd) ≤ x < (mean+2sd)]; 4=x > (mean+2sd). The coded occurrence character values-by-species matrix was then analysed by Correspondence Analysis [19,20]. Euclidean distance was applied to measure dissimilarity on the bases of the coordinates of the first two axes of the correspondence analysis. Distance scores were then clustered to produce phenogram with the unweighted pair-group method using the arithmetic average. In terms of molecular analysis, tissue collection, DNA extraction and sequencing followed the methods reported by Pouyaud *et al.* [13]. Mitochondria of 12SrDNA was selected for the genetic analysis. The method applied Kimura’s distance to produce phylogenetic tree [14].

3. Result and discussion
Twenty four variables were selected for the correspondence analysis. The remaining variables were excluded from the analysis because they do not really contribute much to our understanding of these species. The result showed that a high similarity was found between most species (2.4 <\emph{d}_{eucidean}< 12.8), e.g. *Helicophagus waandersii* and *H. leptorhynchus*. The most different species are *H. typus* and *Pangasius sanitwongsei*. The UPGMA phenogram (Figure 2) revealed six clusters based on linkage.
distance of > 6 grouping the species into the proposed genera or species groups. The first cluster on the top consists of all species belonging to Helicophagus. The second cluster represents Pangasius macronema. The third cluster is Pangasianodon with P. gigas and P. hypophthalmus. Then, there are Pangasius nieuwenhuisii, P. kinabatanganensis, P. humeralis, P. lithostoma, P. nasutus, P. conchophilus, P. sabahensis, P. mekongensis, P. kunyit, P. myanmar, P. krempfi, P. djambal, P. bocourti, P. pangasius in the fourth cluster. The fifth cluster contains three branches, the first branch includes Pteropangasius with two species, P. pleurotaenia, and P. micronemus; the second one consists of Pangasius polyuranodon and P. elongatus; the third consists of P. mahakamensis and P. rheophilus. Finally, the sixth cluster consists of P. larnaudii and P. sanitwongsei.

Figure 2. Phenogram for 28 pangasiid species using UPGMA clustering of similarity coefficient of Euclidean distances derived from 24 variables. Species abbreviation: typ=Helicophagus typus; waa=H. waandersii; lep=H. leptorhynchus; mac=P. macronema; gig=Pangasianodon gigas; hyp=P. hypophthalmus; ple=Pteropangasius pleurotaenia; mic=Pteropangasius micronemus; kin=Pangasius kinabatanganensis; lit=P. lithostoma; hum=P. humeralis; nie=P. nieuwenhuisii; con=P. conchophilus; nas=P. nasutus; sab=P. sabahensis; mek=P. mekongensis; kun=P. kunyit; mya=P. myanmar; kre=P. krempfi; dja=P. djambal; boc=P. bocourti; pol=P. polyuranodon; elo=P. elongatus; mah=P. mahakamensis; rhe=P. rheophilus; pan=P. pangasius; lar=P. larnaudii; san=P. sanitwongsei. Number refer to clusters with a linkage distance of > 6.

Overall, the phenogram found that the analyzed species aggregated into propose genera by Gustiano [1]. However, there are more than one clusters for the Pangasius species due to the extremely characters appearance compare to the others such as the second cluster containing P. macronema is unusual because of the very long maxillary and mandibulary barbel lengths and big eye diameter [17]. The fifth cluster because of the short predorsal length and short mandibular barbel. The sixth cluster (P. larnaudii and P. sanitwongsei) due to the very long palatine tooth plate lengths and a very wide mouth. The remaining Pangasius species (cluster fourth) are distributed in the tree. On this basis, there is a congruency between cluster analysis and the results obtained from previous analysis in the genera and species description [1,14,17].

For encoding 12S rDNA, there were seven hundred and thirty seven nucleotides. Of the 737-bp sequences, 123 sites were polymorphic whereas 72 sites were useful for phylogeny analysis. The phylogenetic dendrogram (Figure 3) showed that the species aggregated to monophyletic group (bootstrap between 77 and 100%) confirming 28 recognised species. The genetic distances among species were between 0.0042 and 0.0561. The most differed is Pangasianodon gigas and P. rheophilus.
(d=0.0561). On the other side, Helicophagus waandersii and H. leptorhynchos (d=0.0042) are the opposite species followed by P. bocourti and P. djambal (d=0.0055), P. nieuwenhuisii and P. humeralis (d=0.0056).

Figure 3. The genetic relationship among pangasiids catfish species using 12S rDNA gene.

Bootstrapping found four monophyletic groups: Pangasianodon hypophthalmus and P. gigas with bootstrap 84% in group 1; Pteropangasius micronemus and P. pleurotaenia with bootstrap 92% in group 2; Helicophagus waandersii, H. leptorhynchos, H. typus with bootstrap 80% in group 3; and P. nieuwenhuisii, P. humeralis, P. polyuranodon, P. macronema with bootstrap 100% in group 4. Besides groups mentioned above, P. lithostoma, P. nasutus, P. larnauidii, P. santwongsei, P. rheophilus, P. conchophilus and P. kinabatanganensis are distributed in the tree. Although some branching have low information on phylogeny, the related species still enable to have high bootstrap value.

Four genera, Helicophagus; Pangasianodon; Pteropangasius; and Pangasius were detected by biometric analysis on 28 species observed in the present study. In the previous study done by Gustiano [1] and Gustiano & Pouyaud [11] reported similar phenomena. According to Gustiano [1], the osteological and genetic data validated four monophyletic genera. The results above mentioned that various analyzed showed congruency for the existing species and the possible genera. The significant incongruence between molecular and morphological-based phylogeny is rare [21]. They are complementary each other’s. Except the genus Pangasius, the members were spread out in different cluster. The molecular analyzed also indicated there were poor resolution of the genus Pangasius. Pouyaud et al. [13] reported the similar result for allozymes and cytochrome-b analyzed. According to [22,23,24] poor resolution in the dendrogram might be caused by strong radiation in short period that enabled to form synapomorphies. However, the related species, such as Helicophagus leptorhynchos and H. waandersii or Pangasius djambal and P. bocourti were very clear.

The phylogeny tree [1] and dendrogram of molecular data, "Pangasius (except for the five endemic species P. kinabatanganensis, P. lithosoma, P. humeralis, P. nieuwenhuisii, and P. myanmar)" is more recent compare to other genera. Using mitochondrial gene, estimated the speciation of Pangasiidae was about 5-8 million years ago (mya) until the Late Pleistocene. A similar case on Hemibagrus nemurus populations in Southeast Asia was also reported by Dodson et al. [25]. Geographical conditions land lock causing different niche and food habits for a long times causing various morphological and taxonomic diversity known at present (specialized groups) as mentioned in the taxonomic revision works. For instance, large vomerine tooth plates are on seven endemic Pangasius species from Borneo (Kalimantan). Very short additional vomeral tooth plates are in P. sabahensis and P. mahakamensis, and combining the vomerine tooth plate is in P. rheophilus. Disappearing the additional vomeral tooth plate is in the remaining four species. Enormous Asian Pangasius from the mainland have a long
additional vomeral tooth plates. Among the species, Genus *Pteropangasius* distribute in all over Southeast Asia. Therefore, before separating the islands in the southern part of Asia about 20 mya, this taxon may already have existed. Model of *Helicophagus typus* divergence and distribution showing vicariance event followed by allopatric distribution. *H. typus* exists in the Barito and Kapuas Rivers (Kalimantan) as well as Musi and Batang Hari Rivers (Sumatera), while *H. waandersii* occurs in the Musi, Batang Hari (Sumatera), and Pahang River (Malay Peninsula). The third species, *H. leptorhynchos* exists in the Mekong and Chao Phraya River. A similar model of distribution for *Hemisilurus* catfishes in the same area was also reported by Bornbusch & Lundberg [26].

The remaining species enabled to be divided into the species Indian subcontinent, Indo-China and the Indonesian Archipelago. The mainland Asian species decreases toward the Indonesian Archipelago and vice versa. The evidence indicating that the remaining species is more recent than *Pteropangasius* and *Helicophagus*, exist after separating of the island from continent in the last glacial event. Allopatric of closely related species is suggesting the vicariant speciation. Examples can be seen in the genus *Pangasius* where *P. conchophilus* occurs in Indo-China and *P. nasutus* in the Sunda Region (the Indonesian archipelago); *P. bocourti* in Indo-China and *P. djambal* in the Sunda Region; the distribution of the group *P. elongatus, P. mahakamensis* and *P. polyuranodon*; the distribution of the group *P. kunyit, P. sabahensis* and *P. mekongensis*. The study found that only *Pangasianodon gigas* and *Pangasius mekongensis* are endemic to the Mekong River while seven species are endemic to Kalimantan/Borneo. The existing of endemic species in Kalimantan e.g. *Pangasius humeralis, P. lithostoma, P. kinabatanganensis, P.nieuwenhuisii, P. rheophilus* and *P. mahakamensis* indicates the area as the centre of speciation due to water barrier or land lock for a long period to be very different ecology for supporting speciation. The species common to Sumatra and Kapuas, and not found in the Mahakam River, are considered to be new, Pleistocene migrants and those also occurring in the Mahakam originate from an older invasion [27].

Study on the tertiary fossil records showed that fish fauna from Sipang, Central Sumatra contain four species belonging to four extinct genera and 12 species in 11 extinct genera [28]. Nine of the 11 extinct genera occur in the Mahakam and Mekong River. Some pangasiid fossil found in Sumatra, Indonesia [28], India [29], and Thailand [30] support the widely dispersal hypothesis of the Pangasiidae occurred since the Tertiary from the Indonesian Archipelago to the Indian subcontinent. Using the cyt- b evolution rate, the divergence among genera in Pangasiidae is around ≈ 11-13 million years ago (mya) [1] corresponds to the Middle and Late Miocene. Roberts & Jumnongthai [30] confirmed with the presence of the fossil genus *Cetopangasius* from Lake Phetchabun in the north-central of Thailand. Based on divergence rate estimates for allozymes, the divergence is about ≈ 19 mya. From the study found, the genus *Pangasius* had high adaptive radiation is ≈ 8-10 mya with cytochrome b and ≈ 9 mya with allozymes. Nowadays, from 28 species exist, only *P. pangasius* is in the Indian sub-continent. The same species occurs in the Irrawadi basin in Myanmar which has one endemic species namely *P. myanmar*. Compared to *Pangasius* species from Southeast Asia, the appearance of Indian subcontinent one (including Myanmar) are a little not the same. However, both are still in the same genus. Pangasiids catfishes evolution showed that adaptive radiation happened after the vicariant speciation. Based on the study on pangasiids catfishes, the existence of the six endemic species from Kalimantan in specific parts of the rivers are a good example of species speciation in Southeast Asia. The results of this study are another illustration of the consistency of freshwater fish fauna distribution and its historical biogeography.

4. Conclusion
The tree revealed that all species analyzed for proposed genera aggregated. Biometric analysis on 28 species supports that they should belong to four genera confirms the results from genetic analysis.

Acknowledgement
The authors thanks to colleagues who work together in the Catfish Project for their help in collecting fish samples. Special thanks due to G. Teugels [late] and L. Pouyaud for his encouragement to the
The present study. The authors also grateful to AARD for the fellowship to complete this study through “PAATP” Project at KU LEUVEN, Belgium. The study was a part of “Catfish Asia” project financed by the European Union (contract IC 18-CT 96-0043).

References
[1] Gustiano R 2003 Taxonomy and phylogeny of Pangasiidae Catfishes from Asia (Ostariophyisis, Siluriformes) (Belgium: KU Leuven, Ph.D. Thesis) p 296.
[2] Legendre M, Pouyaud L, Slembrourck J, Gustiano R, Kristanto A H, Subagia J, Komarudin O, Sudarto, and Maskur 2000 IARD Journal 22 1-16.
[3] Subagia J and Gustiano R 2005 Sainteks 13 8-12.
[4] Vidthayanon C 1993 Taxonomic revision of the catfish family Pangasiidae (Tokyo: Tokyo University of Fisheries, Ph.D Thesis) p 203.
[5] Gustiano R and Pouyaud L 2007 Indonesian Aquaculture Journal 2 107-112.
[6] Gustiano R 2007 Buletin Plasma Nutafah 15 91-100 [in Indonesian].
[7] Pouyaud L and Teugels G G 2000 Ichthyol. Explor. Freshwater 11 193-200.
[8] Pouyaud L, Gustiano R, and Teugels G G 2002 Cybium 26 243-252.
[9] Gustiano R, Teugels G G, and Pouyaud L 2003 J. Nat. Hist. 37 357-376.
[10] Gustiano R, Prakoso V A, Ath-thar M H F, Radona D, and Kusmini I 2019 Indonesian Fisheries Research Journal 25 123-131.
[11] Gustiano R and Pouyaud L 2008 Indonesia Aquaculture Journal 3 13-22.
[12] Gustiano R, Prakoso V A, and Ath-thar M H F 2018 Indonesian Fisheries Research Journal 24 99-115
[13] Pouyaud L, Teugels G G, Gustiano R, and Legendre M 2000 Journal of Fish Biology 56 1509-1538.
[14] Pouyaud L, Gustiano R, and Teugels G G 2004 Indonesian Journal of Agricultural Science 5 45-62.
[15] Karinthanyakit W and Jondeung A 2012 Journal of Fish Biology 80 2549-2570.
[16] Azlina Z M, Daud S K, Siraj S S, Aliabadian M, and Moghaddam F Y 2013 Iran. J. Anim. Biosyst. 9 17-28.
[17] Gustiano R and Pouyaud L 2005 Zuriat 16 66-73.
[18] Gustiano R 2004 Indonesian Journal of Agricultural Science 5 70-74.
[19] Sneath P H A and Sokal R R 1973 Numerical taxonomy: the principles and practices of numeral classification (San Francisco: WH Freeman & Co. San Francisco) p 573.
[20] Manly B F J 2000 Multi variate statistical methods: A primer (Florida: Chapman & Hall/CRC) p 215.
[21] Hillis D M and Wiens J J 2000 Molecules versus morphology in systematics: Conflict, artifact, and misconception. In Phylogenetic analysis of morphological data (Washington: Smithsonian Institution Press) p 19.
[22] Johns G C and Avise J C 1998 Evolution 52 1135-1146.
[23] Martin A P and Bermingham E 1998 Molecular Phylogenetics and Evolution 9 192-203.
[24] Lewin R 1997 Patterns in evolution: the new molecular view (USA: Scientific American Library) p 246.
[25] Dodson J J, Colombani F, and Ng P K L 1995 Molecular Ecology 4 331-346.
[26] Bornbusch A and Lundberg J G 1989 Copeia 2 434-444.
[27] De Beaufort L F 1951 Zoo geography of the land and inland waters (London: Sidwick and Jackson) p 208.
[28] Sanders M 1934 Verhandelingen van het Geologish Mijnbouwkundig Genootschap 11 1-143.
[29] Romer A S 1947 Vertebrate Paleontology (Illionis, USA: The University of Chicago Press) p 687.
[30] Roberts T R and Jumnongthai J 1999 Natural History Bulletin of the Siam Society 47 153-189.