Diploid and tetraploid African *Barbus* (Osteichthyes, Cyprinidae): on the coding of differential gene expression

Patrick Berrebi [1], Christian Lévêque [2], Ghislaine Cattaneo-Berrebi [1], Jean-François Agnèse [1], Jean-François Guégan [3] and Annie Machordom [4]

[1] *Institut des Sciences de l'Evolution (URA 327, CNRS), case 064, Université Montpellier II, Place E. Bataillon, 34095 Montpellier Cedex 05, France.*
[2] *Antenne ORSTOM, Laboratoire d'Ichtyologie Générale et Appliquée, Muséum National d'Histoire Naturelle, 43 rue Cuvier, 75231 Paris Cedex 05, France.*
[3] *Laboratoire de Parasitologie Comparée (URA 698, CNRS), Université Montpellier II, Place E. Bataillon, 34095 Montpellier Cedex 05, France.*
[4] *Museo Nacional de Ciencias Naturales, J. Gutierrez Abascal, 2; 28006 Madrid, Spain.*

Abstract

Three species in the group of "small" *Barbus* from West Africa were analyzed by enzyme electrophoresis to assess their genetic differentiation. Comparison with a species of "large" *Barbus* from the same region showed that the "small" *Barbus* are certainly diploid and the "large" tetraploid. They clearly form two distinct lineages. Phenetic dendrograms (Nei distances) and cladograms (compatibility networks) of the genus *Barbus* are proposed, based on three African diploid species, two diploid species from Saudi Arabia and Southeast Asia, one African tetraploid species, and two French tetraploid species. These trees reveal two sets of species, *i.e.* diploid and tetraploid. Several methods of data processing are suggested for overcoming the difficulties involved in simultaneously analyzing species with different ploidy levels.

Keywords : Tetraploid fish, allozymes, coding methodology, Cyprinidae, African Barbus.

Barbus *africains diploïdes et tétraploïdes: le codage des gènes à expression différentielle.*

Résumé

Trois espèces du groupe des « petits » barbeaux d'Afrique de l'Ouest ont été analysés par électrophorèse des protéines enzymatiques afin d'estimer leur différenciation génétique. La comparaison avec une espèce de « grand » barbeau de la même région a montré que les « petits » barbeaux seraient diploïdes et les « grands » tétraploïdes. Ils constituent certainement deux lignées distinctes. Des arbres phénétiques (distances de Nei) et cladistiques (réseaux de compatibilité) du genre *Barbus* sont proposés comprenant les trois espèces diploïdes africaines, deux espèces diploïdes d'Arabie et d'Asie du sud-est, une espèce tétraploïde africaine et deux espèces tétraploïdes de France. Ces arbres présentent deux ensembles : les diploïdes d'un côté et les tétraploïdes de l'autre. Plusieurs méthodes de traitement des données sont proposées pour résoudre les difficultés dues au traitement simultané d'espèces à niveau de ploïdie différent.

Mots-clés : Poissons tétraploïdes, allozymes, méthodes de codage, Cyprinidae, *Barbus* africains.

INTRODUCTION

The genus *Barbus sensus lato* is currently comprised of the order of 800 species in Africa, Europe, and Asia, many of which show little resemblance to the type species *Barbus barbus* (Linnaeus). In Asia, barbels are usually classified in the genus *Puntius* Hamilton, 1822. *Puntius sensus lato* has been subdivided into *Puntius sensus stricto*, *Capoeta* Valencienne, and *Barbodes* Bleeker, 1859. According to Magton and Arai (1989), "some workers adopt only *Puntius sensus lato* and criticize the classification of three genera (Taki *et al.*, 1978). Others adopt these three genera, but their interrelationships have not been known (Wu *et al.*, 1977)." Other genera and subgenera have also been characterized in eastern Asia (Chu and Kottelat, 1989): *Spinibarbus* Oshima, 1919, *Spinibarbichthys* Oshima, 1926, *Tor* Gray, 1834, *Parator* Wu *et al.*, 1963, *Paraspinibarbus* Chu and Kottelat, 1989, etc. However, as pointed out by Howes (1987), nothing is gained by attributing these species to different genera simply because they populate different continents.

In Africa, it is acknowledged that species now classified in the genus *Barbus* belong to at least two groups, *i. e.* the "large" barbels characterized by their scales with many parallel stria and a dorsal fin with nine to eleven branched rays, and the "small" barbels whose scales have a small number of divergent stria, and whose dorsal fin has seven or eight branched rays. This subdivision has been confirmed by osteological differences (Howes, 1987). Based on a study of the form of the upper jaw, Mahnert and Géry (1977) characterized a third group, which for the moment only contains one species, *Barbus jae*, classified in the "small" barbels. It should be noted that while studying the gill parasites of this species, Birgi and Lambert (1987) observed monogeneans with characters differing from those of specimens infesting other *Barbus*, which confirms the observation of Mahnert and Géry (1977).

In the absence of a comprehensive examination, and considering the many species that have probably not yet been described, it can be estimated that there are about 200 to 250 species of small barbel and 60 to 80 species of large barbel in Africa (Lévêque and Daget, 1984). Several names for genera and subgenera have been proposed by different authors (review in Lévêque and Daget, 1984) but in contrast with Asia, ichthyologists working in Africa have preferred to adopt a conservative position, and the name *Barbus sensus lato* continues to be widely used pending a clearer definition of the phylogenetic relationships between the different groups.

Small and large barbels are widely distributed throughout Africa. However, certain North African species show more morphological similarities with European species than with tropical species. Moreover, in South Africa, Skilton (1985) has reported the following: "The large *Barbus andrewi* and *B. serra* differ in several distinctive ways from other African *Barbus* species but are similar to the European *Barbus barbus*. Whether this similarity is reflective of a closer phylogenetic relationship has not been investigated."

In general, it is difficult to base the systematics of small *Barbus* on morphology in view of the limited literature on comparative studies. For example, it has been shown that the number and size of barbels on the fish can vary depending on the size of the individual or on particular populations of the same species (Lévêque, 1989). The morphological characteristics are generally very similar and the meristic characters are of limited use. Color is an important character, but it also varies with the populations and disappears during fixing procedures. Consequently, long field experience is required for correct species identification.

To develop new criteria that could improve the description of phylogenetic relationships between species and groups of *Barbus*, genetic studies have been carried out on the African species. These species have an important characteristic: the work of Agnèse *et al.* (1990) has shown that the two groups (small and large) have different ploidy levels.

In tetraploid fishes, the number of loci is not twice that of their diploid ancestors. According to Ohno (1970), Ferris and Whitt (1977), and Buth (1983), a tetraploid event is followed by gradual differentiation of homologous loci and homologous chromosomes. The segregation of chromosomes changes from tetrasomic to disomic. In other words, although the species remain tetraploid in terms of the number of chromosomes, it acquires a typically diploid function. This change is accompanied by a large decrease in the number of active loci, which is revealed by enzyme electrophoresis. This process, which Ohno (1970) has called functional diploidization, has been interpreted as an accumulation of deleterious mutations on one of the homologous loci, causing it to be silenced. However, the silencing is not lethal since the second locus maintains enzyme production. According to Ohno (1970), the silencing of certain loci can also result from different regulation in different organs.

Thus, there is clearly a larger number of active loci in tetraploids. This phenomenon creates data processing difficulties in any comparative calculation or analysis of species with different ploidy levels, *i. e.* the difference in the number of markers, hence variables, has to be taken into account.

The Catostomidae, an entirely tetraploid American family, is closely related to the Cyprinidae. Between 1977 and 1984, numerous studies appeared focused on its phylogeny and evolution. A great part of these established phylogenies based on the duplicated or diploidized state of the different loci, taking no account of the allelic divergences (Ferris and Whitt, 1978; Ferris *et al.*, 1979; Buth, 1979 a).

Other studies have compared Catostomidae species using allelic frequencies to compute Nei distances and identities (Nei, 1972). Since there is no variation of the duplicate or diploidized state of a locus when comparing populations of the same species, there was no difficulty in coding data (Ferris et al., 1982; Buth and Crabtree, 1982). It is also the case for very closely related species (Buth, 1977).

However, other studies have established comparisons between species showing variations in the expression of the duplicated genes. Ferris and Whitt (1977 and 1978) provide Nei's distances but do not explain how coding procedures were determined. Buth (1980) established genetic distances between very divergent tetraploid. To explain the coding methology used for the calculation of Nei's distances, this author refered to Engel et al. (1973) who, however, do not propose any clear methodology.

Lastly, Ferris and Whitt (1977) are the only ones to establish genetic distances between diploid and tetraploid species, but provide no explanation.

In summary, the abundant literature treating of phylogeny and evolution of the American tetraploid species do not permit us to know what options were chosen to bypass the difficulties exposed in _figure_ 1.

under the number MNHN 1989-988) from the Pampana (=Jong), a coastal river of Sierra Leone,

- 5 individuals of the species _B. guineensis_ (MHNH 1989-1008 and 1989-1009) from the Konkoure in the Republic of Guinea,

- 6 individuals of the species _B. macrops_ from the Tinkisso, a high affluent of the Niger in the Republic of Guinea.

To situate these species in a wider context, we also analyzed:

- 11 individuals of the species _B. bynni occidentalis_ (MNHM 1989-987), considered to be tetraploid (Agnèse et al. 1990) and belonging to the group of large African _Barbus_, captured in Guinea and Mali,

- 30 _Barbus_ specimens from southern France, _i. e._ 15 _B. barbus_ and 15 _B. meridionalis_, which are known to be tetraploid (Wolf et al., 1969; Sofradzija et al. 1973),

- 14 _individuals of the species B. apoensis_ from the Taïf region (Saudi Arabia), assumed to be diploid,

- 10 diploid individuals (Taki et al., 1977) of the species _B. schwanenfeldi_ originating in Thailand, obtained commercially.
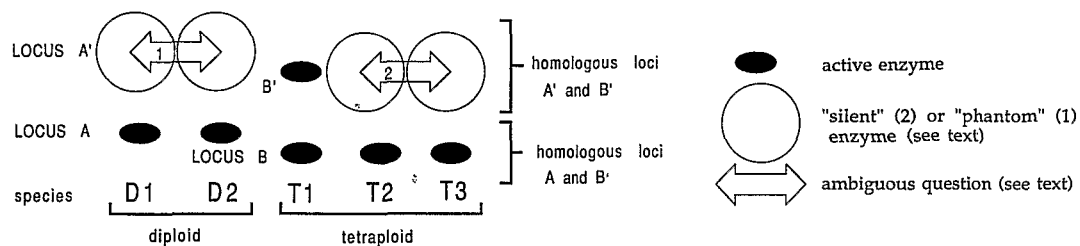

Figure 1. — Diagram illustrating the coding difficulties encountered in phylogenetic studies including both diploid and tetraploid species. Arrow 1: duplication of a locus by tetraploidization. Arrow 2: loss of a locus by functional diploidization in a tetraploid species.

In the present paper, we demonstrate a way of data coding. We analyze several species of small West African _Barbus_ to test homogeneity inside the group. Simultaneous analysis of _Barbus_ living in other regions or continents was carried out to define the phylogenetic relationships of these small _Barbus_ with other diploid (Saudi Arabia and Thailand) and tetraploid _Barbus_ (Africa and Europe).

MATERIALS AND METHODS

Sampling

Nineteen small African _Barbus_ individuals were examined, belonging to three species:

- 8 individuals of the species _B. ablabes_ (deposited in the Museum national d'Histoire naturelle, Paris

Biochemical analysis

All samples were analyzed by starch gel electrophoresis of enzymes. Only muscle and liver were analyzed. The methods cited in Pasteur et al. (1987) were used, as adapted by Berrebi et al. (1988) and Agnèse et al. (1990).

Data coding

The main methodological problem in genetic studies of populations of tetraploid and diploid organisms is the different number of enzyme loci in the two forms, which means that regardless of the analytic method, there is a difference in the number of variables. Although tetraploid species do not express twice the number of enzyme loci, they do express between 30 and 50% more (Woods and Buth, 1984).

As a theoretical example, let us consider 5 species (*fig.* 1). Locus A of the two diploid species D1 and D2, corresponds to loci B and B' of a tetraploid species T1 and to a single locus B in two tetraploid species T2 and T3 (loss of locus B' by functional diploidization).

(i) To equalize the number of loci (and thus the variables) between D species (diploids) and T species (tetraploids), a "phantom" locus can be assigned to the D species, denoted by A', which only would have existed if the diploid species had become tetraploid. This locus is considered to be homozygous for a null allele.

It may seem surprising to use loci which have never existed ("phantom" loci) to code data on diploid species. It is an artificial method to force a comparaison (diploid and tetraploid species) which would be mathematically impossible (different number of variables) but biologically realistic since the diploid species most likely gave rise to the tetraploid ones by a sudden transformation.

To use these methods based on the coding of "phantom" and "silent" loci, one must clearly state the coding conventions.

In this study, all the consequences of the two methods proposed are analysed below by comparing the results obtained in the diploid-tetraploid treatement with that obtained on the diploid species only (see *figures* 3 and 4, opposed to *figure* 5).

(ii) To equalize the number of loci in the tetraploid species, a silent (inactivated) locus can be assigned to species T2 and T3, denoted by B', which is homozygous for a null allele.

The following questions arise. Can phantom locus A' be considered to have the same null allele in the two diploid species, or different null alleles (arrow 1 in figure 1). The same coding problem occurs in the tetraploid species: does silent locus B' have the same null allele in the two species, or different null alleles (arrow 2)? To resolve this problem, two coding methods can be used.

● the "minimizing" method consists of considering that any phantom or silent locus has the same null allele, which is coded in the same way.

● the "maximizing" method consists in considering any null allele of a phantom or silent locus to be different from all the others.

The terms "minimizing" and "maximizing" were chosen because genetic distances (or any other method of quantitative analysis) are affected in these directions by the two coding methods.

The criteria used for determining homologies between loci then have to be specified. Again using the theoretical example in *figure* 1, locus A is arbitrarily considered to be consistently homologous with locus B, which is the least mobile locus in the tetraploid species. There is an exception: if locus A of a diploid has at least one allele with same mobility as

an allele of a tetraploid locus, the two loci (which consequently have a common point) are considered to be homologous, even if it concerns the most rapid locus of the tetraploid species.

Data analysis

All analyses described in this section were done twice using the two coding methods, and the consequences were examined. The elementary results shown in *table* 1 were analyzed in two ways:

● Phenetic analyses were based on allelic frequencies. Nei (1972) distances were calculated and dendrograms were constructed according to the algorithm of the FITCH program in the software package PHYLIP (Felsenstein 1985). This algorithm allows different evolutionary speeds in the same lineage (absence of a molecular clock).

● Cladistic analyses were based on the presence (coded 1) or absence (coded 0) of different alleles. Analyses of compatibility were carried out using the CLIQUE program, which is also in PHYLIP. This analysis determines the phylogenetic system using only characters (alleles) that are compatible (occurring only once, in the form of an appearance or disappearance in the system). For more details, *see* Agnèse *et al.* (1990).

RESULTS

, There are several difficulties involved in interpreting the zymograms of tetraploid species (Buth 1980; Berrebi *et al.* 1988). As an example, one of the most complex systems, Glucose-6-phosphate isomerase (=GPI), is shown in *figure* 2. According to the criteria described above, homologies can be determined between loci that migrate toward the anode and those that migrate toward the cathode. Among the former, locus *Pgi*-1 is remarkably stable, since allele 100 occurs in European tetraploids and certain African, Saudi Arabian, and Asian diploids and so can be considered as ancestral. The homology is not ambiguous.

Loci *Pgi*-2 and *Pgi*-3 code for very mobile electromorphs. Homology is also easy to determine with the locus for rapid electromorphs in the diploids. In the latter, locus *Pgi*-3 is thus a phantom and is considered to be homozygous for a null allele. So, homology can easily be established with the locus having "fast" electromorphs in the diploids, particularly since there is also homology (between the locus 2 of the diploids and the loci 2 and 3 of the tetraploids) in the preferential expression of these loci in the liver.

Lastly, among the enzymes migrating toward the cathode, loci *Pgi*-4 and *Pgi*-5 are only found in tetraploids. For this reason, these loci are phantoms in the diploids.

B. mac.   B. gui.   B. abl.   B. apo.   B. sch.   B. occ.   B. mer.   B. bar.

3/120

3/100   3/100   3/100

2/115   2/115   2/115   2/113

2/110   2/110   2/105   2/110   2/110   2/109

2/090   2/099   2/100   2/100

1/200   1/200   1/200

1/100   1/100   1/100   1/100

1/005   1/005

4/100   4/100   4/100

5/100

●   homodimere molecule
◐   heterodimere molecule
○   echo spot
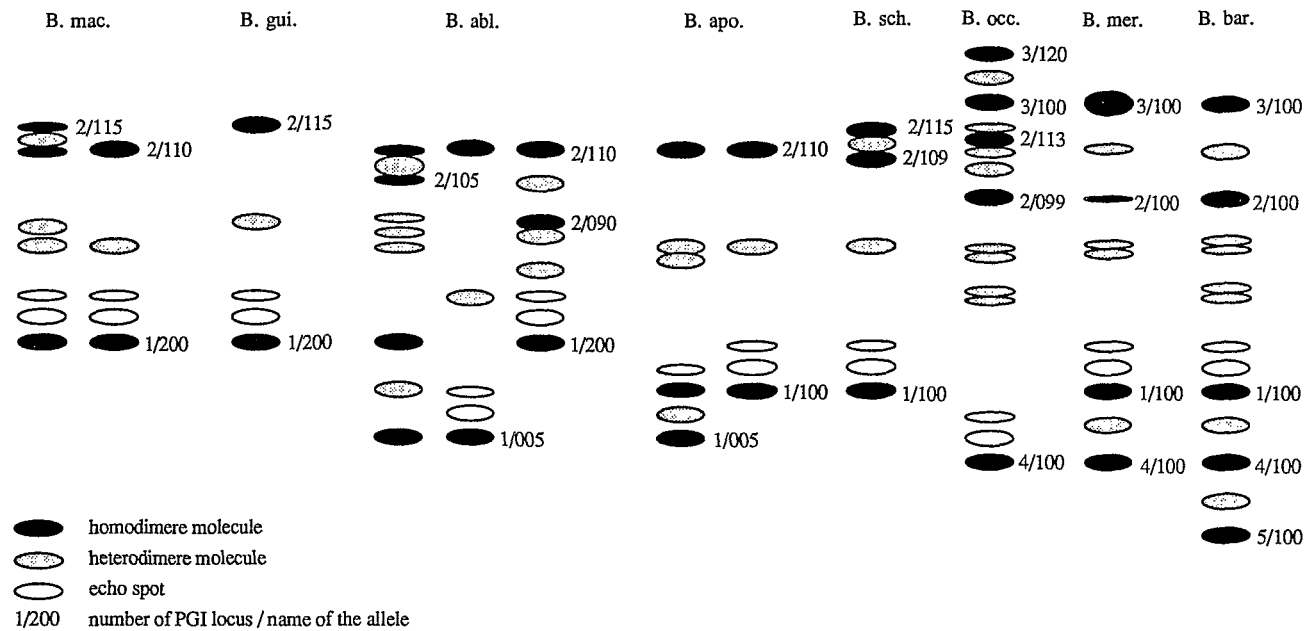1/200   number of PGI locus / name of the allele

Figure 2. — Zymogram of the Glucose-6-phosphate isomerase (GPI) system in the 6 species studied.

The PGI system of *B. barbus* shows the presence of a fifth locus. It does not correspond to a polymorphism of the *Pgi*-4 locus because the pattern show (*figure* 2) is identical in several hundred individuals of that species, from France and Belgium (unpublished data). This fifth locus is considered as a tandem duplication. Even if this phenomenon is considered as rare, it has been mentioned in the case of tetraploid species by Ferris and Whitt (1978), Buth (1979 a, 1980 and 1982) and by Crabtree and Buth (1981).

All the enzyme systems were interpreted using these rules. The general results are summarized in tables of allele frequencies. *Table* 1 was derived using the minimizing code in which all the null alleles of phantom and silent loci were denoted similarly (000 in this case). The results of the maximizing code are not shown. However, at the top of figures 3 to 5, loci *Pgi*-2 and *Pgi*-3 are described with the two types of code. The names of the alleles, except for those of *B. meridionalis*, are not standardized with those of Agnèse *et al.* (1990), since the analyses were not done simultaneously.

The number of active loci on the zymograms varied from 14 to 18 in the species known to be tetraploids (the two French species). Thus, the large African barbel, *B. bynni occidentalis*, also appears to be tetraploid, as noted by Agnèse *et al.* (1990). In contrast, the constant number of ten active loci in the other species designates them as diploids. It should be noted that these figures are approximate, since in the tetraploids it is impossible to distinguish a system with one active locus (thus, with a silent locus) from a

system with two loci that have not diverged, *i.e.* having the same alleles.

All of these analyses, illustrated in *figures* 3 to 5, take into account the diploid/tetraploid division defined above. Each figure comprises: (A) the example showing the coding of the rapid anodic loci of the PGI system, (B) the table of Nei genetic distances, (C) the FITCH dendrogram obtained from this table, and (D) the CLIQUE cladogram based on the presence or absence of alleles.

The following questions can be raised: (i) what is the effect of the two types of coding, and (ii) is it legitimate to analyze species of different ploidy levels?

Figure 3 shows an analysis concerning all the species, regardless of their ploidy level, using the maximizing code. The analysis in *figure* 4 also concerns all the species, but using the minimizing code. *Figure* 5 shows·these analyses on the diploid species treated separately. In this case, different codes are not used since the problems shown in figure 1 do not occur.

It can first be noted that the two types of coding considerably alter most of the genetic distances (*figures* 3 and 4):

● between two similar diploid species, such as *B. macrops* and *B. guineensis* whose Nei distances are 0.145 or 0.991 using the minimizing or maximizing codes, respectively,

● between two dissimilar diploid species, e. g. *B. macrops* and *B. apoensis* whose distances are 0.514 or 2.194,

● between two dissimilar tetraploid species, the coding affected one of the distances (*B. barbus*/*B.*

Table 1. — Allelic frequencies of 21 loci analyzed using the minimizing code. Null alleles (phantom or silent) are denoted by 000*. MACR = *Barbus macrops*; GUIN = *B. guineensis*; ABLA = *B. ablabes*; APOE = *B. apoensis*; SCHW = *B. schwanenfeldi*; OCCI = *B. bynni occidentalis*; MERI = *B. meridionalis* and BARB = *B. barbus*.

| Locus	Allele	MACR	GUIN	ABLA	APOE	SCHW	OCCI	MERI	BARB
Pgi-1	000*	0	0	0	0	0	1	0	0
	005	0	0	0.35	0.64	0	0	0	0
	006	0	0	0	0	0	0	0	0
	100	0	0	0.05	0.36	1	0	1	1
	200	1	1	0.60	0	0	0	0	0
Pgi-2	085	0	0	0.06	0	0	0	0	0
	090	0	0	0.11	0	0	0	0	0
	099	0	0	0	0	0	0.45	0	0
	100	0	0	0	0	0	0	1	1
	105	0	0	0.50	0	0	0	0	0
	109	0	0	0	0	0.05	0	0	0
	110	0.94	0	0.33	1	0	0	0	0
	113	0	0	0	0	0	0.55	0	0
	115	0.06	1	0	0	0.95	0	0	0
Pgi-3	000*	1	1	1	1	1	0	0	0
	100	0	0	0	0	0	0.32	1	1
	105	0	0	0	0	0	0.13	0	0
	110	0	0	0	0	0	0.55	0	0
Pgi-4M	000*	1	1	1	1	1	0	0	0
	100	0	0	0	0	0	1	1	1
Pgi-5	000*	1	1	1	1	1 ·	1	0	1
	100	0	0	0	0	0	0	1	0
Pgm-1	095	0	0	0	0.96	0	0	0	0
	096	0	0	0	0.14	0	0	0	0
	098	0	0	0	0	0	1	0	0
	099	0	0.22	0	0	0	0	0	0
	100	0	0	0	0	0	0	1	1
	102	0	0	0.10	0	0	0	0	0
	110	1	0.78	0.90	0	0	0	0	0
	111	0	0	0	0	0.10	0	0	0
	120	0	0	0	0	0.90	0	0	0
Pgm-2	000*	1	1	1	1	1	0	0	1
	090	0	0	0	0	0	1	0	0
	100	0	0	0	0	0	0	1	0
Ldh-1	100	0	0.25	0.86	1	0	0.88	1	1
	130	0	0	0	0	0	0.12	0	0
	135	1	0.75	0.14	0	1	0	0	0
Ldh-2	100	1	1	1	1	1	1	1	1
Mdh-2F	100	0	0	0	0	0	0	1	1
	120	0	0	0	0	0	1	0	0
	125	1	1	1	0	0	0	0	0
	150	0	0	0	0	1	0	0	0
	160	0	0	0	1	0	0	0	0
Mdh-4F	000*	1	1	1	1	1	0	0	1
	095	0	0	0	0	0	1	0	0
	100	0	0	0	0	0	0	1	0
Idh-M1	000*	1	1	1	1	1	1	0	1
	100	0	0	0	0	0	0	1	0
Idh-M2	099	0	0	0	1	0	0	0	0
	100	1	1	0.83	0	0	1	1	1
	150	0	0	0.17	0	0	0	0	0
	160	0	0	0	0	1	0	0	0
Sod-1	100	0	0	0	0	0	1	1	1
	105	0	0.71	0	0	1	0	0	0
	150	1	0.29	1	0	0	0	0	0
	195	0	0	0	1	0	0	0	0
Sod-2	000*	1	1	1	1	1	0	0	1
	095	0	0	0	0	0	1	0	0
	100	0	0	0	0	0	0	1	0
Ak-1	000*	1	1	1	1	1	0	1	0
	100	0	0	0	0	0	0	0	1
	200	0	0	0	0	0	1	0	0
Ak-2	050	1	0	0.40	0	0	0.30	0	0
	099	0	0	0	1	0	0	0	0
	100	0	0.95	0.60	0	0	0.70	1	1
	120	0	0.06	0	0	0	0	0	0
	130	0	0	0	0	1	0	0	0
Aat-1	100	0	0	0	0	0	0	1	1
	125	0.12	0	0.65	1	1	1	0	0
	130	0.88	1	0.35	0	0	0	0	0
Aat-2	000*	1	1	1	1	1	0	0	0
	080	0	0	0	0	0	1	0	0
	090	0	0	0	0	0	0	0	1
	100	0	0	0	0	0	0	1	0
Nb. Active loci		10	10	10	10	10	16	18	14
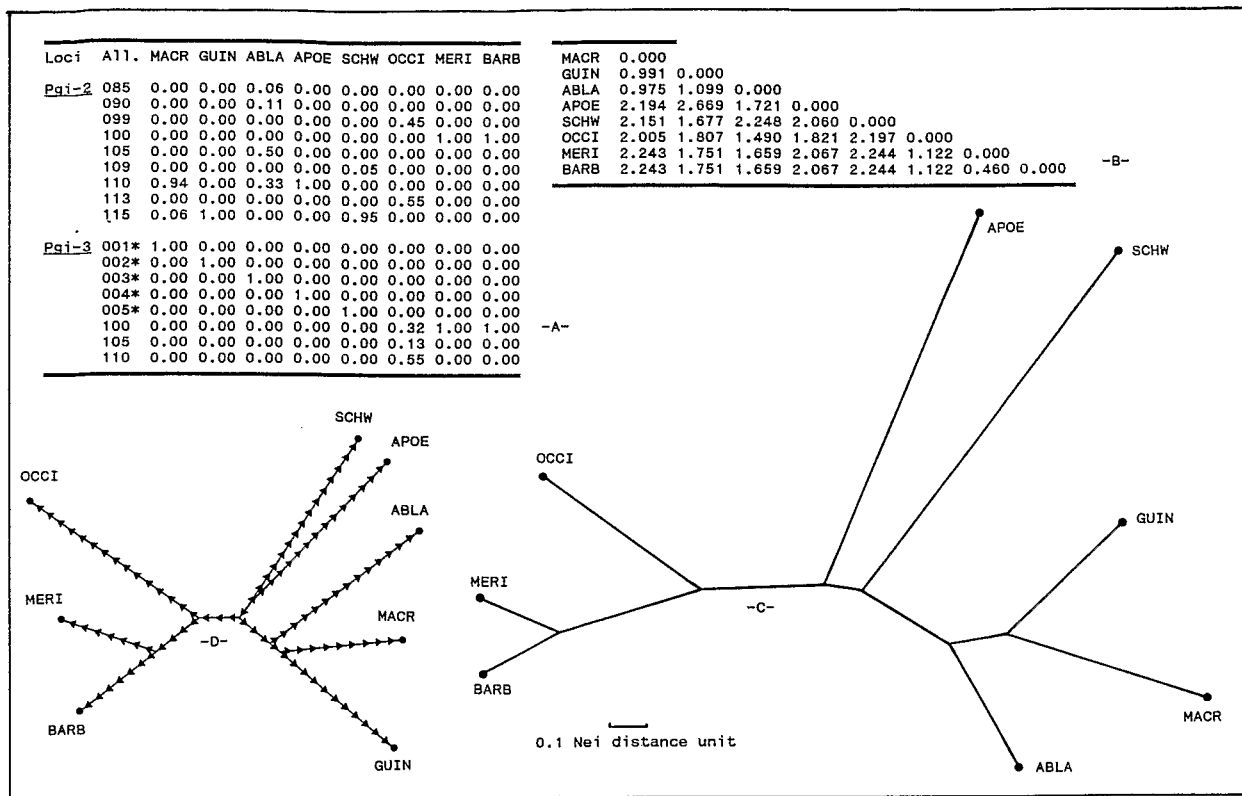
Figure 3. — Results using the maximizing code. A: table of the allelic frequencies of loci *Pgi-1* and *Pgi-2*. B: table of Nei genetic distances of the species considered. C: dendrogram of Nei distances obtained by the FITCH algorithm. D: cladogram obtained using the CLIQUE program (Felsenstein 1985).
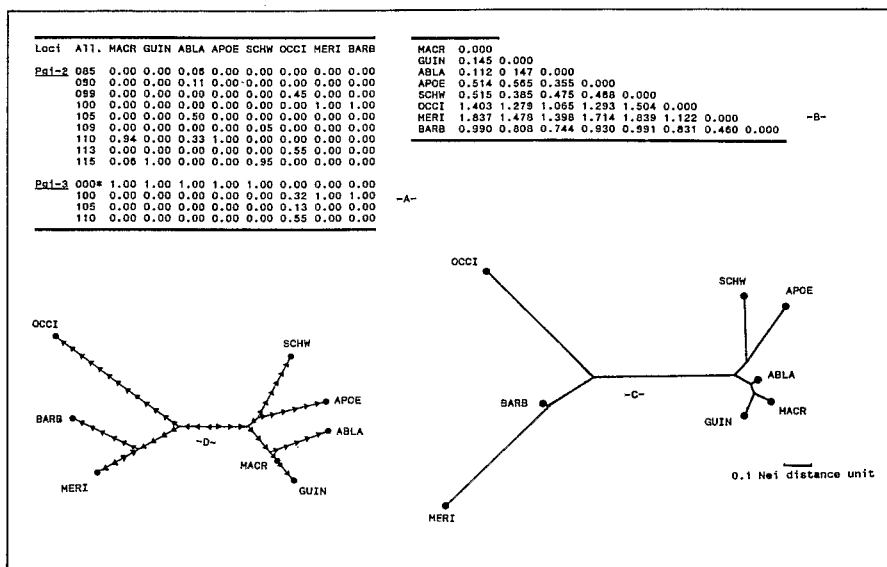

Figure 4. — Results using the minimizing code. *See* legend of figure 3.

bynni occidentalis: 0.831 and 1.122) but not the other (*B. meridionalis/B. bynni occidentalis:* 1.122);

● on the other hand, between two similar tetraploid species (*B. barbus* and *B. meridionalis*) the distance did not change (0.460),

● between diploids and tetraploids (mean value) it varied from 1.285 to 2.099.

The dendrograms obtained using the algorithm of the FITCH program show that coding affects the length of the branches, but not their relative position.
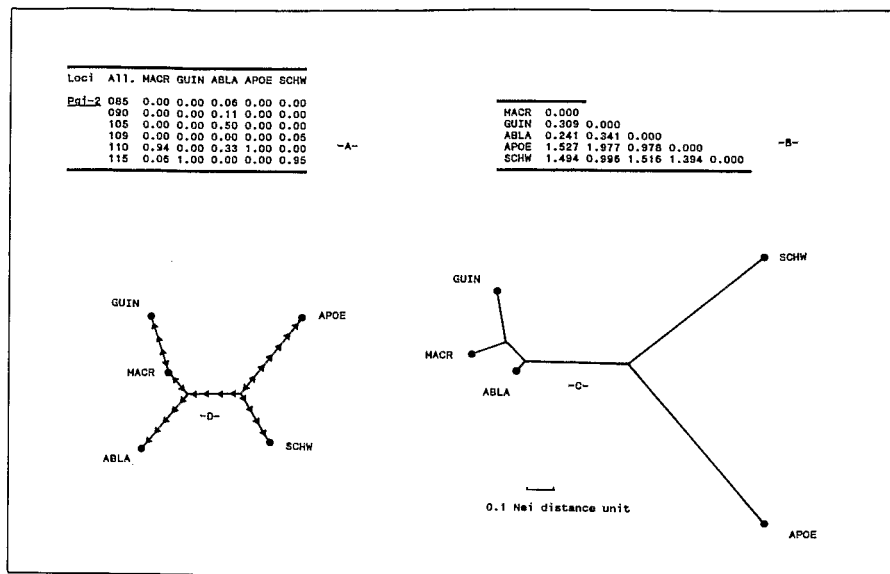
Figure 5. — Results concerning only diploid species. *See* legend of figure 3.

The diploid species form three independent branches, *i. e.* West African, Saudi Arabian, and Southeast Asian species.

In contrast, the two types of coding affect the mean distance between diploids and tetraploids. Paradoxically, comparing this distance with that separating two dissimilar diploid species, it is 2.5 times greater with the minimizing code and the same value with the maximizing code. In other words, using the minimizing code, the species are clearly grouped according to their ploidy level, and the distance between groups is greater than the distance between species, whereas with the maximizing code, the distance between groups is of the same order or less than the distance between species.

Agnèse *et al.* (1990) used the minimizing code. The resulting Nei distances are very similar to those obtained in the present study, except for the distance between the two French *Barbus* species. This is due to the fact that the enzyme systems analyzed were not all identical in the two studies.

The same conclusions can be drawn from a comparison of the CLIQUE networks. The minimizing code more effectively individualizes each ploidal group (6 events separate them, *i. e.* 3 appearances of characters in each group) than the maximizing code (3 events, all of them appearances in the tetraploids). In general, the minimizing code amplifies the distinction between ploidal groups and the maximizing code amplifies the distinction between species. To help in choosing between the methods, these results can be compared with those obtained independently in each ploidal group.

● *Diploids.* — *Figure* 5 shows that the genetic distances (which are true genetic distances since there is no interference from the tetraploids) are still between the minimizing and maximizing values, but are 4 times closer to the distances calculated with the minimizing code.

● *Tetraploids.* — The duality of the coding is inescapable. However, the effect of coding is much smaller than in the diploids. Moreover, it is advisable to consider two null alleles of the same silent locus as two different alleles, which implies using the maximizing code.

● *The distance between diploids and tetraploids.* — This varies from single to double (between 1.285 and 2.099) and there is no basis for choosing either value.

Lastly, we recommend that quantitative phylogenies (genetic distances) be established, including diploids and tetraploids, in a composite manner. This means first establishing the phylogenies of the two ploidal groups, using conventional means with diploids, and the maximizing code with tetraploids. They should then be related by a distance whose value is between those obtained by the two coding methods. This large uncertainty is related to the logical impossibility of comparing species with different ploidy levels.

We realize that this coding method is artificial. However, it is one of the few methods using allozymes, which allows us to place species having different level of ploidy on the same trees. An article in preparation will propose and test some other coding methods. Ambiguity could be avoided using methods that analyze the genetic material without involving ploidy levels, such as ribosomal RNA sequencing or in certain cases by analyzing the polymorphism of the restriction sites of mitochondrial DNA.

With regard to the cladograms (CLIQUE), although the number of events differs greatly depending on the type of coding, the form of the systems is

identical, which is the main result required in this type of analysis.

DISCUSSION

In terms of systematics, the present results show that the species of *Barbus* studied here form two groups with certainly different ploidy levels. With the minimizing code, the two groups are quite distinct, whereas with the maximizing code, the distance (phenetic) or difference in the number of events (cladistics) between groups often have values that are lower than the corresponding values between species. Be that as it may, the two ploidal groups are always independent and no analysis affects this division. If the classification prevailing in Africa is extrapolated, it can be said that there is a division between large (tetraploid) and small (diploid) *Barbus*.

In this study, there are three radiating branches within the group of diploids: West African, Saudi Arabian, and Southeast Asian species (*figure* 5). These species are of course not representative of all diploid *Barbus*, but there is equal distance between the three entities, and no two geographical groups can be shown to be in opposition with the third.

From a technical viewpoint, Iskandar and Bonhomme (1984) have shown that studies concerning species that are dissimilar (spatially, and consequently temporally) entail many problems with regard to the homology of electromorphs. Two distortions can appear when the difference in the mobility of two allozymes is too small to be detectable;

• either the proteins are very similar, which is a clear sign of relatedness, and a significant evaluation error is not introduced by considering them to be identical;

• or the proteins are very different, and their resemblance on the gel arises from the fact that the resulting charge is the same by coincidence. However, since this coincidence is governed by chance, such errors are distributed homogeneously in all analyses. Because of this, it can be expected that the results will be falsified in the form of a "technical" decrease in distances. We estimate that this should not affect the overall structure of the phylogenetic trees, except when the branch of a lineage in a cladogram depends on a single marker, which never occurs here, since nearly 50% of the active loci in African diploid species have at least one allele in common with each of the species of the Middle East and Asia.

According to hypotheses proposed by various authors, such as Darlington (1948 and 1957), Banarescu (1973) and Almaça (1976), the genus *Barbus* originated from Asia and dispersed in two ways.

1) an Asia-Europe dispersion, the "Siberian branch", which would have occurred between the

Oligocene and Pliocene. All the European species studied until now (about a third) are tetraploid (Wolf et al. 1969, Sofradzija and Berberovic 1973, Hafez 1981, Triantaphyllidis et al. 1981, Berrebi et al. 1988, etc.).

2) An Asia-Africa dispersion, forming what has been called the "Ethiopian branch". There are no cytogenetic studies in the literature. The present study and that of Agnèse et al. (1990) show that the African *Barbus* may be polyphyletic, each phylum having a different ploidy level:

(i) In the tetraploid species, the morphology of large *Barbus*, such as *B. parawaldroni*, described by Lévêque et al. (1987), is in every way comparable to the Moroccan species of the subgenus *Labeobarbus*, whose morphology has been summarized by Berrebi (1981). We thus consider that this African tetraploid lineage extends to Morocco (which has previously been proposed by Boulenger 1919) and belongs to the "Ethiopian branch" of Darlington.

(ii) The diploid species of small Barbus appear to form a distinct lineage, which may have colonized Africa in a different period. From the present results, it cannot be assumed that *B. apoensis* (Saudi Arabia) constitutes a forerunner of this colonization. This hypothesis would have to be based on a comparison of African species with many Asian diploid species, particularly those classified in the genus *Puntius sensu lato*, as well as with diploid species of the Middle East occupying an intermediate geographic position.

CONCLUSION

The results reported above shed light on the phylogeny of the genus in West Africa. However, the conclusions largely depend on the ploidy level of the species concerned. The European species are well-documented in this respect, but the number of chromosomes in the African species remains to be clearly determined.

The use of concepts of African colonization is essentially practical. No data yet confirm the Asia-Africa direction of the dispersion of this genus. Moreover, the overall aspect of this hypothesis has been the subject of debate. Géry (1969) and Roberts (1969) consider that the cyprinids are of African origin and that the dispersion was toward Asia, among other areas. Novacek and Marshall (1976) situate the origin in South America, whereas Briggs (1979) returns to the idea of a Southeast Asian origin. It should lastly be noted that although the movement of species between continents depends on geological possibilities, there have probably also been two-way migrations, with species dispersing simultaneously in all possible directions.

REFERENCES

Agnèse J. F., P. Berrebi, C. Lévêque, J. F. Guégan, 1990. Two lineages, diploid and tetraploid, demonstrated in African species of *Barbus* (Osteichtyes, Cyprinidae). *Aquat. Living Resour.*, 3, 305-311.

Almaça C., 1976. La spéciation chez les Cyprinidés de la Péninsule Ibérique. *Rev. Trav. Inst. Pêches marit.*, **40**, 399-411.

Banarescu P., 1973. Origin and affinities of the freshwater fish fauna of Europe. *Ichthyologia*, **5**, 1-8.

Berrebi P., 1981. Contribution à l'étude du sous-genre *Labeobarbus* (genre *Barbus*, Poissons, Cyprinidés) au Maroc. *Bull. Inst. Sc. Rabat.*, **5**, 59-72.

Berrebi P., G. Lamy, G. Cattaneo-Berrebi, J. F. Renno, 1988. Variabilité génétique de *Barbus meridionalis* Risso (Cyprinidae): une espèce quasi monomorphe. *Bull. Fr. Pêche Piscic.*, **310**, 77-84.

Birgi E., A. Lambert, 1987. Les Dactilogyridae (Plathelminthes, Monogenea) des Cyprinidae du genre *Barbus* au sud Cameroun. *Rev. Hydrobiol. trop.*, **20**, 37-48.

Boulenger G. A., 1919. La distribution en Afrique des Barbeaux du sous-genre *Labeobarbus*. *C. R. hebd. Séanc. Acad. Sc., Paris*, **169**, 1016-1018.

Briggs J. C., 1979. Ostariophysian zoogeography: an alternative hypothesis. *Copeia*, **1**, 111-118.

Buth D. G., 1977. Biochemical identification of *Moxostoma rhothoecum* and *M. hamiltoni*. *Biochem. Syst. Ecol.*, **5**, 57-60.

—, 1979 *a*. Creatine Kinase variability in *Moxostoma macrolepidotum* (Cypriniformes; Catostomidae). *Copeia*, **1**, 152-154.

—, 1979 *b*. Duplicate gene expression in tetraploid fishes of the tribe Moxostomatini (Cypriniformes, Catostomidae). *Comp. Biochem. Physiol.*, **63**, 7-12.

—, 1980. Evolutionary genetics and systematic relationships in the Catostomid genus *Hypentelium*. *Copeia*, **2**, 280-290.

—, 1982. Glucosephosphate-isomerase expression in a tetraploid fish, *Moxostoma lachneri* (Cypriniformes, Catostomidae): evidence for retetraploidization? *Genetica*, **57**, 171-175.

—, 1983. Duplicate isozyme loci in fishes: Origins, distribution, phyletic consequences and locus nomenclature. *In*: Isozymes: Current Topics in Biological and Medical Research, Vol. 10, M. C. Rattazzi, J. G. Scandalios, G. S. Whitt eds., A. R. Liss, New York, 381-400.

Buth D. G., C. B. Crabtree, 1982. Genetic variability and population structure of *Catostomus santaanae* in the Santa Clara drainage. *Copeia*, **2**, 439-444.

Chu X. L., M. Kottelat, 1989. *Paraspinibarbus*, a new genus of Cyprinid fishes from the Red River Bassin. *Jpn. J. Ichtyol.*, **36**, 1-5.

Crabtree C. B., D. G. Buth, 1981. Gene duplication and diploidization in tetraploid Catostomid fishes *Catostomus fumeiventris* and *C. santaanae*. *Copeia*, **3**, 705-708.

Darlington P. J., 1948. The geographical distribution of cold-blooded vertebrates. *Rev. Biol.*, **23**, 105-123.

—, 1957. Zoogeography: The geographical distribution of animals. Wiley Ed., London.

Engel W., J. Schmidtke, W. Vogel, U. Wold, 1973. Genetic polymorphism of lactate dehydrogenase isoenzymes in the carp (*Cyprinus carpio*) apparently due to a "null allele". *Biochem. Genet.*, **8**, 281-289.

Felsenstein J., 1985. "PHYLIP" Phylogeny Inference Package, 2.8 version. MS DOS/IBM PC Microcomputer version, Univ. Washington, Seattle.

Ferris S. D., G. S. Whitt, 1977. Loss of duplicate gene expression after polyploidisation. *Nature*, **265**, 258-260.

—, 1978. Phylogeny of tetraploid Catostomid fishes based on the loss of duplicate gene expression. *Syst. Zool.*, **2**, 189-205.

Ferris S. D., D. G. Buth, G. S. Whitt, 1982. Substantial Genetic Differentiation among Populations of *Catostomus plebeius*. *Copia*, **2**, 444-449.

Ferris S. D., S. L. Portnoy, G. S. Whitt, 1979. The roles of speciation and divergence time in the loss of duplicate gene expression. *Theor. Popul. Biol.*, **1**, 114-139.

Ferris S. D., D. G. Buth G. S. Whitt, 1982. Substantial Genetic Differentiation among Populations of *Catostomus plebeius*. *Copeia*, **2**, 444-449.

Géry J., 1969. – The freshwater fishes of South America, pp. 828-848. *In*: Biogeography and ecology in South America, E. J. Fittkau *et al.* eds., W. Junk, The Hague.

Hafez R., 1981. Études cytogénétiques et évolution génétique chez quelques Cyprinidés de la région Midi-Pyrénées. Thèse dr., Inst. Nat. Polytechnique Toulouse, 362 p.

Howes G. J., 1987. The phylogenetic position of the Yugoslavian Cyprinid fish genus *Aulopyge* Heckel, 1841, with an appraisal of the genus *Barbus* Cuvier and Cloquet 1816, and the subfamily Cyprinidae. *Bull. Br. Mus. Nat. Hist., Zoology*, **52**, 165-196.

Iskander D., F. Bonhomme, 1984. Variabilité électrophorétique totale à 11 loci structuraux chez les rongeurs muridés (Muridae, Rodentia). *Can. J. Genet. Cytol.*, **26**, 622-627.

Lévêque C., 1989. Remarques taxinomiques sur quelques petits *Barbus* (Pisces, Cyprinidae) d'Afrique de l'Ouest (1re partie). *Cybium*, **13**, 165-180.

Lévêque C., Daget, 1984. Cyprinidae, pp. 217-342. *In*: Check list of the freshwater fishes of Africa. Daget J., Gosse J. P., Thys van den Audenaerde eds., MRAC-ORSTOM, 410 p.

Lévêque C., D. F. E. Thys Van Den Audenaerde, K. Traoré, 1987. Description de *Barbus parawaldroni* sp. n. (Pisces, Cyprinidae) d'Afrique occidentale. *Cybium*, **11**, 347-355.

Magtoon W., R. Arai, 1989. Kariotypes of five *Puntius* species and one *Cyclocheilichthys* species (Pisces, Cyprinidae) from Thailand *Bull. Nat. Sci. Mus. Tokio*, **15**, 167-175.

Mahnert V., J. Géry, 1977. La machoire supérieure de certains "*Barbus*' s. l. (Pisces, Cypriniformes) africains. *Rev. Suisse Zool.*, **84**, 229-236.

Nei M., 1972. Genetic distance between populations. *Am. Nat.*, **106**, 283-292.

Novacek M. J., L. G. Marshall, 1976. Early biogeographic history of Ostariophysian fishes. *Copeia*, **1**, 1-12.

Ohno S., 1970. Evolution by gene duplication, Springer-Verlag, New York.

Pasteur N., G. Pasteur, F. Bonhomme, J. Catalan, J. Britton-Davidian, 1987. Manuel de génétique par électrophorèse des protéines. Lavoisier, Paris, 217 p.

Roberts T., 1969. Osteology and relationships of characoid fishes particularly the genera *Hepsetrus*, *Salminus*,

Hoplias, *Cenolucius* and *Acestrorhynchus*. *Proc. Calif. Acad. Sci.*, **36**, 391-500.

Skelton P. M., 1985. Distribution patterns and biogeography of non tropical southern African freshwater fishes. *Paleoecol. Africa*, **17**, 211-230.

Sofradzija A., L. Berberovic, 1973. The chromosome number of *Barbus meridionalis petenyi* Heckel (Cyprinidae, Pisces). *Bull. Sci. Acad. RSF Yougosl.*, **18**, 77-78.

Taki Y., A. Katsuyama, T. Urushido, 1978. Comparative morphology and interspecific relationships of the Cyprinid genus *Puntius*. *Jpn. J. Ichtyol.*, **25**, 1-8.

Taki Y., T. Urushido, A. Suzuki, C. Serizawa, 1977. A comparative chromosome study of *Puntius* (Cyprinidae: Pisces). I. South-east Asian species. *Proc. Japan Acad.*, **53**, 6 231-235.

Triantaphyllidis C. D., H. Damianakis, P. S. Economidis, J. Karakousis, 1981. Genetic variation in greek barbel populations. I. Esterases, LDH, MDH, ME and PGM in *Barbus meridionalis* (Pisces, Cyprinidae). *Comp. Biochem. Physiol.*, **70**, 278-293.

Wolf U., H. Ritter, N. Atkin, S. Ohno, 1969. Polyploidization in the family Cyprinidae, order Cypriniformes. I. DNA content and chromosome set in various species of Cyprinidae. *Humangenetik*, **7**, 240-244.

Woods T. D., D. G. Buth, 1984. High level of gene silencing in the tetraploid Goldfish. *Biochem. Syst. Ecol.*, **12**, 415-421.