Genetic differentiation in *Scriptaphyosemion geryi* (Lambert, 1958) suggests the existence of a species complex

by

Jean-François AGNÈSE* (1), Christian CAUVET (2) & Raymond ROMAND (3)

Abstract. – *Scriptaphyosemion geryi* (Lambert, 1958) is a well-defined taxa easily distinguished from the other congeneric species by a red marbled pattern, unpaired fins with deep blue blotches, blue margin and red submargin. A zigzag mid longitudinal dark line on sides characterizes females. For long, due to a high phenotypic variability, this taxa was suspected to encompass several cryptic species. We sampled 20 populations representing the entire species range and characterized the genotypes (cytochrome *b*, mitochondrial DNA) of their specimens. Results allowed grouping populations in five clusters composed of geographically-linked populations. Compared with those observed in other species from the subgenus *Chromaphyosemion* (Radda, 1971), all these results strongly suggested that *S. geryi* encompassed five different species.

Received: 18 Apr. 2012
Accepted: 19 Apr. 2013
Editor: A. Gilles

Key words
Nothobranchiidae
Aplocheilidae
Scriptaphyosemion geryi
West Africa
Killifish
Speciation

Résumé. – La différenciation génétique chez *Scriptaphyosemion geryi* (Lambert, 1958) révèle l'existence d'un complexe d'espèces.

Scriptaphyosemion geryi est un taxon bien défini dont les mâles se distinguent facilement des mâles des autres espèces du genre par un aspect général rouge marbré, des nageoires impaires avec des taches bleu profond, une bande marginale bleue et une bande submarginale rouge. Les femelles sont caractérisées par une ligne longitudinale noire en zigzag sur les côtés. Depuis longtemps on pense, en raison de la grande variabilité phénotypique observée, que ce taxon pourrait en fait englober plusieurs espèces cryptiques. Nous avons échantillonné 20 populations réparties sur toute l'aire de répartition de l'espèce et représentant toute la variabilité phénotypique connue. Nous avons caractérisé chaque population génétiquement (étude des séquences du cytochrome *b*). Les résultats ont permis de regrouper les populations en cinq groupes géographiquement ordonnés. La comparaison de l'étendue de la différenciation génétique chez *S. geryi* avec celle que l'on observe chez d'autres groupes d'espèces du sous genre *Chromaphyosemion* (Radda, 1971) suggère fortement que *S. geryi* est un taxon qui recouvre en fait cinq espèces différentes.

Scriptaphyosemion Radda & Pürzl, 1987 are Aplocheilidae fishes that inhabit rainforest rivulets in tropical West Africa from Senegal to Liberia (Huber, 2006). A conspicuous feature of these fishes is their distinct sexual dimorphism. Males are brightly coloured, slightly larger than females and possess enlarged unpaired fins, which are not present in the cryptically coloured females. The taxonomy of the genus is mainly based on differences in the colour patterns of the males, since the meristic and morphometric characters normally used in ichthyology do not allow to delineate species in this taxon (Amiet, 1987; Scheel, 1974, 1990). The genus is composed of 12 species and *S. geryi* (Lambert, 1958) is distinguished from the other congeneric taxa by the presence of red dots or stripes in the caudal fin like *S. cauveti* (Romand & Ozouf, 1995) but differing from this later species by a red pigmentation of the body not dominant over the background *versus* dominant in *S. cauveti*. *Scriptaphyosemion geryi* has two remarkable features. On one hand, since their discovery, many different natural populations have been collected from Senegal to Sierra Leone (Wildekamp and Van der Zee, 2003). This large range is quite uncommon in *Scriptaphyosemion* (as in many Aplocheilidae fishes). Species usually occupy a smaller range mainly due to important speciation events based on phenotypic differentiation rather than on ecological differentiation. This pattern of evolution implies that all these species share the same ecological niche and subsequently are barely found in sympatry. On the other hand, the different natural populations of *S. geryi* are phenotypically differentiated (Wildekamp and Van der Zee, 2003). Although no real phenotypic analysis has been made, one can observe that populations from Senegal, Guinea Bissau and the North-West part of Guinea are characterized by specimens with numerous dots and a large submarginal red band in the anal fin. Populations from coastal rivers of Western Guinea are characterized by specimens with a greenish anal fin with a thin submarginal red band and few or even no red dots. Specimens from populations around Conakry (except population from Coyah) are characterized by an anal fin with

(1) Département conservation et domestication, Institut des sciences de l'évolution (ISE-M) UMR 226 UM2-CNRS-IRD, Case courrier 63, Université de Montpellier II, 34095 Montpellier CEDEX 05, France.
(2) 10 allée René Cuiry, 76420 Bihorel, France. [christian.cauvet@wanadoo.fr]
(3) Institut de génétique et de biologie moléculaire et cellulaire, Université de Strasbourg, 1 rue Laurent Fries, BP 10142, Illkirch CEDEX, France. [romand@igbmc.fr]
* Corresponding author [jean-francois.agnese@ird.fr]

vertical red bands. Specimens from coastal rivers South to Conakry to Sierra Leone are characterized by an anal fin entirely red with few green dots. Specimens from the Coyah population are easily distinguished from all the other populations because they have only red dots, versus red stripes on their body.

This large range, together with this phenotypic variability, indicate that the taxon *S. geryi* could represent different cryptic species. To test this hypothesis, we sampled different natural populations representing all the known phenotypic variability and genetically characterised these samples. If *S. geryi* represents a species complex, a large amount of genetic differentiation (compared to what is observed between other Aplocheilidae species) is expected together with a congruence with phenotypic differentiation.

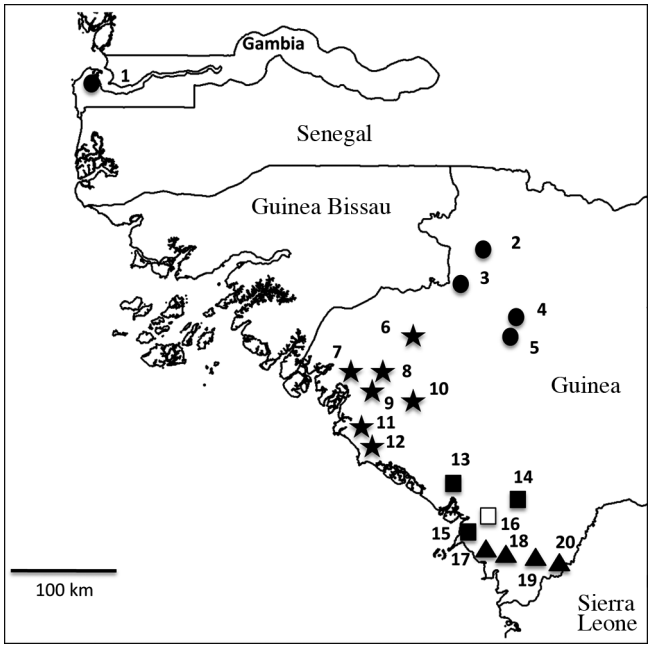

Figure 1. - Map showing sample distribution. Population numbers referred to table I. The different symbols used refer to population clustering as detailed in the text: ●, clade 1; ★, clade 2; ▲, clade 3; ■, clade 4; □, clade 5.

MATERIALS AND METHODS

Sample collection

More than two hundred individuals were collected from 19 locations across the distribution range of *S. geryi* in Guinea (Tab. I; Fig. 1). Specimens for DNA analysis were anaesthetized with phenoxyethanol and then preserved in 90% alcohol. Additional specimens (population 1) were obtained from hobbyists and included in the analyses (Tab. I). Total sampling included the most marginal populations of the species, the north-western in Gambia (population 1) and the south-eastern in Guinea (the most eastern population known coming from Port Loko in Sierra Leone which is about 50 km of population 20). Voucher specimens for natural populations have been registered in the laboratory (ISE-M, Montpellier) under references indicated in table I.

Table I. - Population number (No), Localities, coordinates, voucher references and GenBank accession numbers.

	Coordinates				
No	North	West	Locality	Voucher reference	GenBank No.
1	13.23.749	16.38.696	Abuko	ISEM-JFA-GAB	JX044136
2	11.59.858	13.31.202	Dombradji	ISEM-JFA-G15	JX044120
3	11.42.217	13.44.660	Déola	ISEM-JFA-G16	JX044121
4	11.30.528	13.15.209	Sud Gaoual	ISEM-JFA-G12	JX044119
5	11.14.795	13.10.422	Kakoni	ISEM-JFA-G10	JX044118
6	11.11.746	14.04.196	Tiangi	ISEM-JFA-G17	JX044122
7	10.42.836	14.28.905	Kamsar	ISEM-JFA-G18	JX044123
8	10.42.100	14.23.018	Kolaboui	ISEM-JFA-G19	JX044124
9	10.33.646	14.26.780	Bintimodia	ISEM-JFA-G20	JX044125
10	10.42.411	13.50.598	Wondiré	ISEM-JFA-GCH	JX044117
11	10.27.822	14.24.994	Mankountan	ISEM-JFA-G21	JX044126
12	10.19.059	14.20.887	Kinkon	ISEM-JFA-G22	JX044127
13	10.03.960	13.43.195	Tarene	ISEM-JFA-G24	JX044129
14	9.50.458	13.01.368	Kamara Bounyi	ISEM-JFA-G02	JX044128
15	9.40.351	13.30.082	Dubreka	ISEM-JFA-G26	JX044130
16	9.41.094	13.21.818	Coyah	ISEM-JFA-G01	JX044116
17	9.26.617	13.18.316	Maferenya	ISEM-JFA-G30	JX044134
18	9.26.457	13.12.375	Dandayah	ISEM-JFA-G27	JX044131
19	9.14.284	12.57.685	Farmoreya	ISEM-JFA-G28	JX044132
20	9.12.929	12.54.199	Malifu	ISEM-JFA-G29	JX044133

Laboratory protocols

Total genomic DNA was extracted from fin tissues using the protocol described in Sambrook *et al.* (1989). One fragment of cytochrome *b*, from the mitochondrial genome, was used to reconstruct a mitochondrial gene tree and to evaluate genetic divergence. Standard PCR was performed with two specific primers (Ctb-F1 5'AACCACCGTTGTTATTCAAC3' forward and ctb-R1 5'CTCCCAAAGCCAGAATTCTAAA3' reverse). The amplification protocol consisted of 35 cycles beginning with 3 min at 93°C for initial denaturation followed by cycles of 30 sec at 93°C, 30 sec at 53°C for annealing, 1 min 30 sec at 72°C for extension, with a final 5 min extension step at 72°C.

Phylogenetic analyses

All sequences were edited and aligned using Seqscape version 2.5 (Life Technology) and

subsequently inspected manually. *S. cauveti* (Romand and Ozouf, 1995) and *S. guignardi* (Romand, 1981) were selected as out-group because they are related to *S. geryi* (Huber, 2006).

Prior to analyse the sequence data, we used jModeltest version 0.1.1 (Posada, 2008) to conduct hierarchical likelihood ratio tests (hLRT) to determine the best DNA substitution model of nucleotide for the cytochrome *b* fragment. Model and values obtained were used for subsequent analyses.

Tree search analyses were performed under Bayesian inferences using MrBayes (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003; Altekar *et al*., 2004), Maximum Likelihood using PhyML 3.0 (Guindon *et al*., 2010) and Parsimony using TNT (Goloboff, 1999; Nixon, 1999; Goloboff *et al*., 2008).

In order to appreciate the importance of genetic diversity in *S. geryi*, comparisons have been done with differentiation observed between species of the genus *Aphyosemion*, subgenus *Chromaphyosemion* for which cytochrome *b* sequences have been deposited in GenBank (Sonnenberg, 2007)

Aphyosemion ecucuense (Sonnenberg, 2007) (GenBank number EU249508.1), *A. erythron* (Sonnenberg, 2007) (EU249500.1), *A. melanogaster* (Legros, Zentz & Agnèse, 2005) (EU249513.1), *A. splendopleure* (Brüning, 1929) (EU056941.1), *A. kouamense* Legros, 1999 (EU249497.1), *A. malumbresi* Legros & Zentz, 2006 (EU249501.1), *A. punctulatum* (Legros Zentz & Agnèse, 2005) (EU056937.1), *A. poliaki* Amiet, 1991 (EU056935.1), *A. loennbergii* (Boulenger, 1903) (EU056929.1), *A. lugens* Amiet, 1991 (EU056930.1), *A. bitaeniatum* (Ahl, 1924) (DQ522281.1), *A. bivittatum* (Lönnberg, 1895) (DQ522279.1), *A. riggenbachi* (Ahl, 1924) (DQ342222.1).

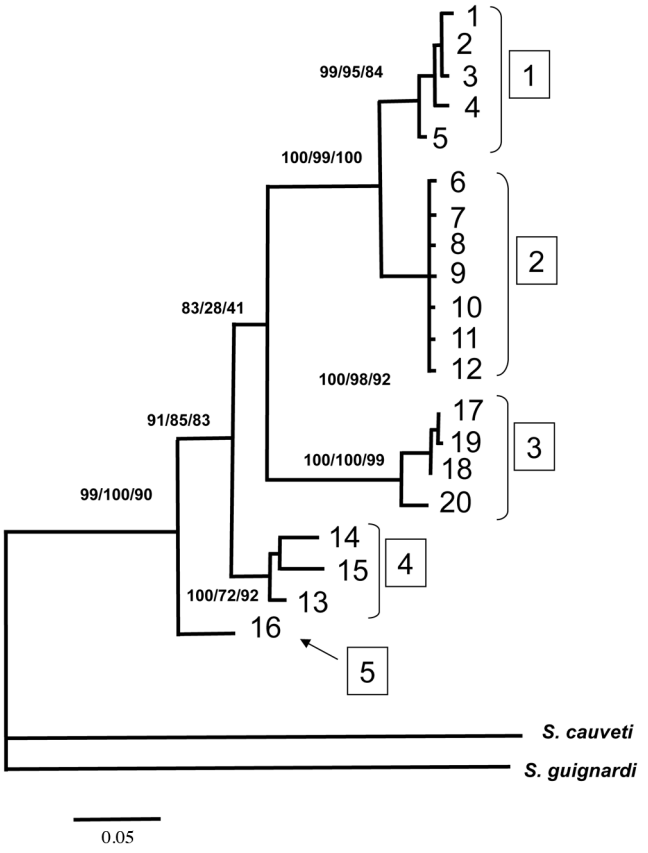

Figure 2. - Consensus tree based on Maximum Likelihood (PHYML), Parsimony (TNT), and Bayesian phylogenetic analysis (number of generation = 6 10^8, burn-in = 25%). Numbers above or below the branches are percentages of bootstrap values based on 1000 replicates for respectively each method (except for Bayesian inferences where the numbers represent the percentage of trees with best likelihood scores in which the particular node was found). Numbers on the right side of the tree represent genetically-defined clades.

RESULTS

Sequence diversity

A 798 base-pair alignment for cytochrome *b* region from the mitochondrial genome was obtained after trimming the ends of each sequence.

A total of 27 specimens were sequenced, including the rooting species (GenBank accession numbers are presented in table I for *S. geryi* specimens; JX044135 for *S. guignardi* and JX044137 for *S. cauveti*). When two specimens of the same population were sequenced (populations 3, 7, 10, 13 and 16), they were always characterized by an identical haplotype. No common haplotype was observed between two different populations.

A total of 122 variable sites were identified for all samples (excluding the outgroup), from which 87 were parsimony informative (i.e. shared by at least two different sequences). Using the corrected Akaike Information Criterion (AIC) in jModeltest (Posada, 2008), the optimal model of sequence evolution was HKY + I + G. Base frequencies estimated from the data were A = 0.2756, C = 0.2727, G = 0.1402 and T = 0.3116; proportion of invariant sites was I = 0.454; and gamma distribution shape parameter was C = 0.487. These criteria were used for subsequent analyses.

The Bayesian phylogenetic analysis (number of generation = 600 millions, burn-in = 25%) recovered five well-supported (83-100% of posterior probabilities) major clades (1-5; Fig. 2). The same clades were also highly supported by the two other analyses, Maximum Likelihood (PHYML) and Parsimony (TNT) for which bootstrap supports calculated from 1000 replicates (Felsenstein, 1985) were respectively between 84% and 99% or 72% and 100%.

The first clade was composed of haplotypes found in populations 1 to 5 while the second clade was composed of haplotypes from populations 6 to 12. These two well dif-

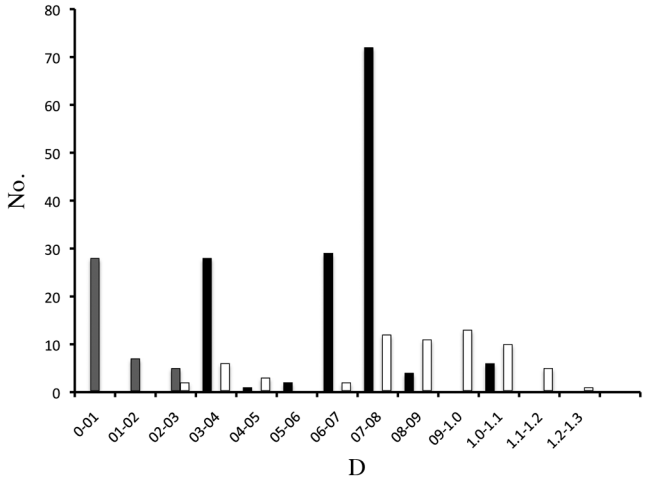

Figure 3. - Histogram representing the distribution of the Kimura 2 distances (D) observed between the 13 different *Chromaphyosemion* species, in white (ranged from 0.024 to 0.120), between *S. geryi* populations of the same clade, in grey (from 0.001 to 0.026), and between *S. geryi* populations from different clades, in black (from 0.030 to 0.083).

ferentiated clades were grouped together. This assemblage was highly supported by all the analyses (100% of posterior probabilities for MrBayes, 99% and 100% of bootstrap value for TNT and PHYML respectively).

The third well-differentiated clade (100%, 100%, 99%) was composed of haplotypes from populations 17 to 20 and was grouped together with clade 1 and 2 but only the Bayesian analysis (83% of posterior probabilities) supported this clustering.

The two other clades, 4 (haplotypes from populations 13 to 15) and 5 (haplotype from population 16) occupied a basal position relatively to the root of the tree (*S. cauveti* and *S. guignardi*) the last being the more basal.

To appreciate the level of genetic differentiation between the different populations or between the different groups, Kimura 2 distances (Kimura, 1980) obtained were compared with the same distance values that characterized genetic differentiation between 13 different *Chromaphyosemion* species. Figure 3 represents the distribution of the three different series of results obtained. Distances observed between the 13 different *Chromaphyosemion* species were ranged from 0.024 to 0.120 (average value = 0.080). Distances observe between *S. geryi* populations of the same group varied from 0.001 to 0.026 (average value = 0.008). Distances observed between *S. geryi* populations from different groups were ranged from 0.030 to 0.083 (average value = 0.064).

Distance values observed between two *S. geryi* populations of the same group (maximum = 0.026) never exceed values observed between two populations of different groups (minimum = 0.030). This confirmed that each genetic cluster was composed of closely genetically related populations.

When compared with values observed between *Chromaphyosemion* species, inter-group values in *S. geryi*, even smaller on average (0.064 versus 0.080) appeared to be comparable. Some interspecific distances in *Chromaphyosemion* (0.024 between *A. malumbresi* and *A. ecucuense*) being smaller than the minimum inter-group distance observed in *S. geryi* : 0.030 between population 3, clade 1, and population 12, clade 2.

DISCUSSION

Scriptaphyosemion geryi populations possessed different mtDNA haplotypes that allowed their clustering in five genetic clades. Populations 1 to 5 have been clustered in clade 1, populations 6 to 12 in clade 2, populations 17 to 20 in clade 3, populations 13 to 15 in clade 4 and population 16, the sister group of all the other populations could be considered as clade 5.

Compared with genetic differentiation observed in *Chromaphyosemion* species, genetic divergences between the different clades of *S. geryi* appeared to be comparable with what has been observed at the species level. When looking at the geographical distribution of these different clades (except for clade 5 for which only one sample has been studied) it clearly appeared than the different populations belonging to each clade were spatially linked. Clade 1 was spread from the western part of Guinea to Gambia, clade 2 occupied the northern coastal zone of Guinea, clade 3 the southern coastal zone of Guinea, clade 4 the Conakry region. Population 16 (clade 5) was in between clades 3 and 4.

Although no real phenotypic analysis has been made, one can observe some congruence between clades defined on a genetic basis and the different phenotypes 1) from Senegal, Guinea Bissau and the North-West part of Guinea, 2) from coastal rivers of Western Guinea, 3) from the surrounding of Conakry, 4) from coastal rivers South to Conakry to Sierra Leone and 5) from Coyah.

All these different results strongly suggested that *S. geryi* represents a complex of at least five different species. Then one can ask at which taxon or at which clade the *S. geryi* holotype can be associated. *Scriptaphyosemion geryi* has been described in 1958 by Lambert as a subspecies of *Aphyosemion guineense* Daget, 1954 [now *Archiaphyosemion guineense* (Daget, 1954)]. Lambert (1958) description was based on fishes captured in the Conakry-Dubreka region in 1955 that correspond to population 15, clade 4, of the present study which is only 7 km away from Dubreka. Taking into account that the 1954 fishing point is not precisely known, we cannot *a priori* exclude that these fishes could correspond to population 16 (clade 5) less than 20 km away from Dubreka. Lambert (1958) described living males with red dot on their body making some broken lines and with many

dots on anal, caudal and ventral fins. Brown dots forming lines have also been observed in females. These descriptions can fit with phenotypes observed with fishes of clade 4 but do not correspond to fishes from clade 5. Then the Taxon *S. geryi* will have to be restricted to fishes or populations from clade 4.

Acknowledgements. – Authors wish to thank Hassimiou Tall and Sekou Camara from CNSHB (Centre National des Sciences Halieutiques de Boussoura) for their help in facilitating field research. Authors wish also to thank J. Laird for help in fish catching, L. Chirio and P. Polichnowski for providing some living fishes. Sequences newly produced during this work were financed by I.R.D. and obtained through the technical facilities of the "Montpellier Environnement Biodiversité" Research Federation.

REFERENCES

ALTEKAR G., DWARKADAS S., HUELSENBECK J.P. & RONQUIST F., 2004. - Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics*, 20, 3 : 407-415.

AMIET J.L., 1987. - Le genre *Aphyosemion* Myers (Pices, Teleostei, Cyprinodontiformes). Faune du Cameroun, 262 p., 68 pl. Tome 2, Sciences Nat. Compiègne.

FELSENSTEIN J., 1985. - Confidence limits on phylogenies: an approach using bootstrap. *Evolution*, 39: 783-791.

GOLOBOFF P.A., 1999. - Analyzing large data sets in reasonable times: solutions for composite optima. *Cladistics*, 15(4): 415-428.

GOLOBOFF P., FARRIS S. & NIXON K., 2008. - TNT, a free program for phylogenetic analysis. *Cladistics*, 24: 774-786 Available: http://www.zmuc.dk/public/phylogeny/ TNT/ (Program and documentation).

GUINDON S., DUFAYARD J.F., LEFORT V., ASIMOVA M., HORDIJK W. & GASCUEL O., 2010. - New Algorithms And Methods To Estimate Maximum-Likelihood Phylogenies: Assessing The Performance of PhyML 3.0. *Syst. Biol.*, 59(3): 307-332.

HUBER J.H., 2006. - Killi-data catalogue. 1028 p. Paris: Killi-Data Editions.

HUELSENBECK J.P. & RONQUIST F., 2001. - MRBAYES: Bayesian inference of phylogeny. *Bioinformatics,* 17: 754-755.

KIMURA M., 1980. - A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucléotide sequences. *J. Mol. Evol.*, 16 : 111-120.

LAMBERT J.G., 1958. - Poissons Siluriformes et Cyprinodontiformes récoltés en Guinée française, avec la description d'une nouvelle espèce de *Microsynodontis*. *Rev. Zool. Bot. Afr.*, 57(1-2): 39-56.

NIXON K.C., 1999. - The parsimony ratchet, a new method for rapid parsimony analysis. *Cladistics*, 15: 407-414.

POSADA D., 2008. - jModelTest: Phylogenetic Model Averaging. *Mol. Biol. Evol.*, 25: 1253-1256.

RONQUIST F. & HUELSENBECK J.P., 2003. - MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12): 1572-1574.

SAMBROOK J., FRITSCH E.F. & MANIATIS T., 1989. - Molecular cloning: a laboratory manual. 1659 p. New York: Cold Spring Harbor Laboratory Press.

SCHEEL J.J., 1974. - Rivulins Studies. *Ann. Mus. R. Afr. Cent.*, Tervuren, ser IN-8. Sci. Zool., 211: 47-98.

SCHEEL J.J., 1990. - Atlas of killifishes of the old world. 448 p. Neptune City, NJ: TFH Publication.

SONNENBERG R., 2007. - Two new species of *Chromaphyosemion* (Cyprinodontiformes: Nothobranchiidae) from the coastal plain of Equatorial Guinea. *Ichthyol. Explor. Freshw.*, 18: 359-373.

WILDEKAMP R.H. & VAN DER ZEE J.R., 2003. - Aplocheilidae. *In*: The Fresh and Brackish Water Fishes of West Africa, Vol. II (Paugy D., Lévêque C. & Teugels G.G., eds), pp. 325-442. Coll. Faune et Flore tropicales 40. Musée Royal de l'Afrique Centrale, Tervuren, Belgique, Muséum National d'Histoire Naturalle, Paris, France and Institut de Recherche pour le Développement, Paris, France.