



A novel molecular marker for the study of Neotropical cichlid phylogeny

T.M.C. Fabrin^{1,2}, L.S. Gasques³, S.M.A.P. Prioli^{2,4} and A.J. Prioli²

¹Programa de Pós-Graduação em Ecologia de Ambientes Aquáticos Continentais, Universidade Estadual de Maringá, Maringá, PR, Brasil

²Núcleo de Pesquisa em Limnologia, Ictiologia e Aquicultura, Universidade Estadual de Maringá, Maringá, PR, Brasil

³Instituto de Ciências Biológicas, Médicas e da Saúde, Universidade Paranaense, Umuarama, PR, Brasil

⁴Departamento de Biotecnologia, Genética e Biologia Celular, Universidade Estadual de Maringá, Maringá, PR, Brasil

Corresponding author: T.M.C. Fabrin

E-mail: thomazmcf@gmail.com

Genet. Mol. Res. 14 (4): 18131-18139 (2015)

Received August 7, 2015

Accepted October 7, 2015

Published December 22, 2015

DOI <http://dx.doi.org/10.4238/2015.December.22.39>

ABSTRACT. The use of molecular markers has contributed to phylogeny and to the reconstruction of species' evolutionary history. Each region of the genome has different evolution rates, which may or may not identify phylogenetic signal at different levels. Therefore, it is important to assess new molecular markers that can be used for phylogenetic reconstruction. Regions that may be associated with species characteristics and are subject to selective pressure, such as opsin genes, which encode proteins related to the visual system and are widely expressed by Cichlidae family members, are interesting. Our aim was to identify a new nuclear molecular marker that could establish the phylogeny of Neotropical cichlids and is potentially correlated with the visual system. We used Bayesian inference and maximum likelihood analysis to support the use of the nuclear opsin *LWS* gene in the phylogeny of eight Neotropical cichlid species. Their use concatenated to the mitochondrial gene *COI* was also tested. The *LWS*

gene fragment comprised the exon 2-4 region, including the introns. The *LWS* gene provided good support for both analyses up to the genus level, distinguishing the studied species, and when concatenated to the *COI* gene, there was a good support up to the species level. Another benefit of utilizing this region, is that some polymorphisms are associated with changes in spectral properties of the LWS opsin protein, which constitutes the visual pigment that absorbs red light. Thus, utilization of this gene as a molecular marker to study the phylogeny of Neotropical cichlids is promising.

Key words: *LWS*; opsin; Molecular marker; *Cichla*; Nuclear marker

INTRODUCTION

The comparison of data obtained from phylogenetic studies, from which similarities and differences among the homologous characteristics of different species may be inferred, aims to detect phylogenetic evidences of the species (Amorim, 2002). Physical characteristics of an organism and their molecular structure may be employed to reconstitute such relationships.

Despite morphological characteristics being evident and easy to analyze, they may adequately show evolutionary traits of species. However, these are frequently subject to environmental influence, as shown by the formation of different phenotypes. On the other hand, although molecular structures are a direct representation of genetic characteristics (Buso, 2005), they do not always provide sufficient differences to outline the evolution of the studied species (Hillis, 1987). In certain cases, morphological data are not consistent with molecular data (Wiley et al., 1998; López-Fernández et al., 2005). Ancestry outlined by molecular data is referred to as molecular phylogeny (Patwardhan et al., 2014).

Since the analysis of molecular data does not take into consideration the environmental factors that affect morphological characteristics, these have been reliable in ichthyology, especially in phylogeny, systematics, genetic variability, management of species, analysis of pedigree and endogamy, and particularly in the identification of hybrid specimens (Gasques et al., 2013). However, the sequences used in such analyses should demonstrate a rate of evolution and be sufficiently preserved and different to be used as specific markers. Molecular markers, nonetheless, do not always show a direct relationship with physiological or behavioral characteristics that are related to species' evolution (van Hazel et al., 2006).

The identification of new markers that can be used in the study of cichlids and vertebrates, which have variable rates of evolution (Beheregaray, 2008; Macrander, 2010; Gasques et al., 2015), and that are directly related to the behavior or morphological characteristics subjected to selective pressures, is highly relevant (van Hazel et al., 2006).

The use of opsin genes has been suggested in species with well-developed visual systems, as these DNA sequences are directly linked to selective pressure in animals (Schwanzara, 1967). The products of these genes are directly related to visual pigments in vertebrates (Trezise and Collin, 2005). Their expression, and the existence of certain polymorphisms, may be directly related to the ecological niche and thus affect the choice of habitat, prey-predator relationship, and behavior of the cichlid cohort (Smith and Carleton, 2010; Nagai et al., 2011), which have a great variety of colorings and may vary geographically within the same species (Maan and Sefc, 2013).

Since Neotropical cichlids have a well-developed visual system, we suggest that the *LWS* (Long Wavelength Sensitive) gene, which is a member of the opsin family, is able to distinguish between species of this group. This hypothesis is based on the fact that the opsin family gene *SWS1* (Short Wavelength Sensitive) has been shown to be a good marker for inferring phylogeny in vertebrates (van Hazel et al., 2006).

The current analysis aimed to identify a nuclear molecular marker that is able to establish the phylogeny of Neotropical cichlids, and that has a relationship through the visual system.

MATERIAL AND METHODS

PCR and DNA extraction

Eight Neotropical cichlid species were used in the current assay. Individuals of *Cichla piquiti* (N = 4) and *Cichla kelberi* (N = 4) were collected from the Tocantins-Araguaia River basin, while those from *Cichla monoculus* (N = 4) were retrieved from the Solimões River. Individuals of *Crenicichla britskii* (N = 4), *Crenicichla haroldoi* (N = 3), *Astronotus crassipinnis* (N = 3), *Satanoperca pappaterra* (N = 3), and *Geophagus proximus* (N = 3) were collected from the flood plain of the high Paraná River within the Long Term Ecology Research (PELD-CNPq) project.

Samples were stored in 96% ethanol and deposited in the tissue bank of the laboratory of genetics at the Research Nucleus in Limnology, Ichthyology, and Aquiculture (Nupélia) of the Universidade Estadual de Maringá. DNA was extracted using a kit obtained from Promega Wizard Genomics. Previously described primers were employed to amplify the opsin gene *LWS* from exon 1 to 6 (Weadick et al., 2012). The full gene was approximately 2.3 kb in size, and a 874-bp fragment was used, which ranged from exons 2 to 5 and included introns. The PCR program consisted of an initial 3-min denaturation (94°C), followed by 35 cycles of 30 s denaturation (94°C), 30 s annealing (51°C) and 2 min 30 s extension (72°C), followed by a 7-min final extension (72°C).

Amplified fragments were purified as previously described (Rosenthal et al., 1993) and samples for sequencing were prepared with a BigDye Terminator kit, following the manufacturer instructions. Sequences were obtained using an automatic sequencer (ABI3730) and gene sequences of subunit I of the cytochrome oxidase (*COI*) of the Neotropical cichlid species and *LWS* of the *Oreochromis niloticus* were obtained from GenBank (Table 1).

Table 1. Studied species and GenBank accession number.

Species	LWS	COI
<i>Cichla kelberi</i>	Present study	JN988797
<i>Cichla piquiti</i>	Present study	JN988799
<i>Cichla monoculus</i>	Present study	JN988800
<i>Crenicichla britskii</i>	Present study	GU701938
<i>Crenicichla haroldoi</i>	Present study	JN988829
<i>Geophagus proximus</i>	Present study	GU701786
<i>Satanoperca pappaterra</i>	Present study	JN989214
<i>Astronotus crassipinnis</i>	Present study	GU701862
<i>Oreochromis niloticus</i>	AF247128	KC789551

Phylogenetic analysis

Sequences were aligned in Mega 6 (Tamura et al., 2013) from the Clustal W algorithm

(Thompson et al., 1994) and selection of the best nucleotide substitution model was performed by the program jModelTest (Darriba et al., 2012). Two types of phylogenetic reconstruction were performed: the first reconstruction used only the *LWS* gene, whilst the second used a concatenation of the genes *LWS* and *COI* following the model GTR+I+G, from the maximum likelihood statistical methods and Bayesian inference, both of which were compared later. The African cichlid species *Oreochromis niloticus* was used as an outgroup for the analyses. Resulting trees were edited by FigTree.

Maximum likelihood analyses were performed by the program raxMLGUI (Silvestro and Michalak, 2012), and the first analysis used only the sequence of the *LWS* opsin gene. The second analysis utilized the *COI* gene sequence linked to the *LWS* gene sequence. The two genes were independent partitions. In both analyses, rapid bootstrap algorithm was used with 1000 re-samplings available in raxML.

Bayesian inference analyses were performed by the program BEAST (Drummond et al., 2012): the first considered the *LWS* gene sequence alone; the second considered *COI* and *LWS* concatenated as independent partitions. The number of generations was sufficient to obtain effective sample size (ESS) rates of >100, with a relaxed molecular clock.

RESULTS

Sequence of *LWS* ranged throughout the extension, from the start of exon 2 to the end of exon 4, including introns (874 bp), whereas the *COI* gene sequence was 570 bp long. Following concatenation (*COI*+*LWS*), the sequences had a total length of 1444 bp. Table 2 shows 124 polymorphic nucleotide sites in the segment of the sequenced *LWS* gene obtained from the species of Neotropical cichlids under analysis.

Table 3 shows the models selected for each sequence and the characterization of each segment. *LWS* gene sequences seem to be species-specific, although they show only scant diversity among species of the same genus. On the other hand, *COI* was highly variable and is therefore an important marker for use in species identification, as employed in the Barcoding project (Ratnasingham and Hebert, 2013).

Figure 1 shows the phylogenetic reconstruction generated using the *LWS* gene in different statistical methods. Bootstrap rates obtained by the maximum likelihood method provided good support, with more than 70% up to the genus level.

Phylogeny determined by the Bayesian inference method was similar to that inferred by the maximum likelihood method, except for the positioning of species belonging to the genus *Cichla*. Phylogenies failed to reveal any great differences among the statistical methods and the generated phylogeny of Neotropical cichlids was similar to that reported by López-Fenández (2010).

Figure 2 shows phylogenies obtained from concatenated sequences of *COI* and *LWS*, following Bayesian inference (A) and maximum likelihood (B).

In the case of concatenated sequences, although the maximum likelihood method provided good support for all branches (>70%), the Bayesian inference method also showed a good basis for phylogeny. Lowest support was obtained for the genus branches *Geophagus* and *Crenicichla*. Phylogeny inferred by the maximum likelihood method for all concatenated markers was also close to that suggested by López-Fernández (2010).

Table 3. Characterization of molecular markers and selected models of nucleotide substitution.

Marker	LWS	COI
Fragment length	874 bp	570 bp
Substitution model	HKY + G	HKY + I + G
Frequency A	0.2335	0.2514
Frequency T	0.3368	0.2902
Frequency C	0.2097	0.3249
Frequency G	0.2200	0.1335
Gamma shape	0.4900	0.6280
Invariable sites (%)	85.81	62.28
Polymorphic sites (%)	14.19	37.72

HKY + G = Hasekawa, Kishino and Yano; G = gamma distribution; I = invariable sites.

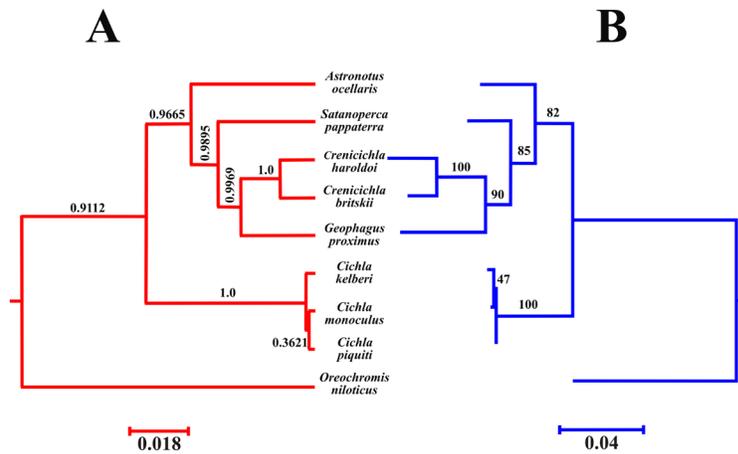


Figure 1. Phylogenetic reconstruction of the *LWS* gene based on the statistical methods of Bayesian inference (A) and maximum likelihood (B).

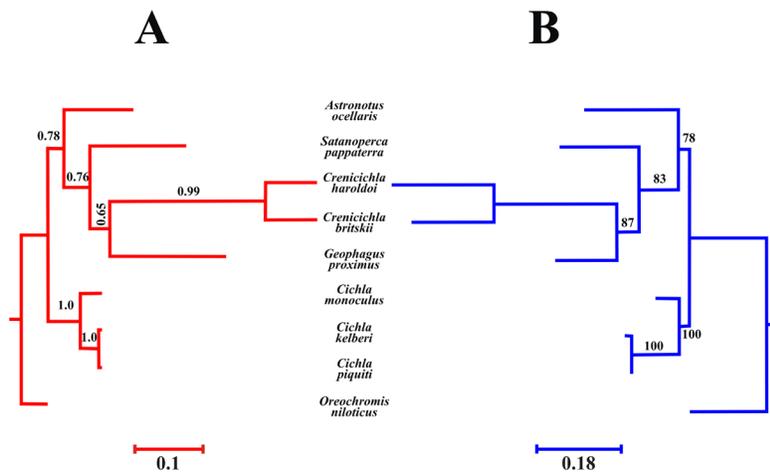


Figure 2. Phylogenetic reconstruction of the *LWS* and *COI* concatenated genes based on the statistical methods of Bayesian inference (A) and maximum likelihood (B).

DISCUSSION

Molecular markers that are associated with a type of selection involved in the speciation process are relevant, since they facilitate phylogeny studies, depending on the species, and the study making it potentially more efficient, because, in general, these regions are under environmental pressure (van Hazel et al., 2006; Fabrin et al., 2014). Since the product of a specific gene may bring about a specific behavior or characteristic, a gene may affect the speciation process. Despite this, several genes may be related to the process of speciation in cichlids (Brawand et al., 2014).

A coherent phylogeny can be obtained through analysis of the opsin *LWS* gene sequence, as reported by López-Fernández et al. (2010). Low variability within the species analyzed may indicate a high degree of phylogenetic information for the sequence (Sullivan et al., 1999; Steel and Penny, 2000). Consequently, the *LWS* gene seems to be a promising marker for Neotropical cichlid species, with good support up to the genus level provided by high bootstrap rates and a high *a posteriori* probability for Bayesian analysis. Miyagi et al. (2012) showed that the *LWS* gene seems to possess species-specific alleles, a fact that also contributes to its use as a marker that can assist with phylogenetic inferences.

Since there is low *LWS* support for species distinction of the same genus, its association with information sequences of greater interspecies variability, such as that of mitochondrial genes, seems adequate. Concatenated sequences of the *COI* and *LWS* genes revealed a greater resolution in species distinction and increased the distance between branches given in phylogeny and its bootstrap and *a posteriori* values.

Although mitochondrial genes seem to be inefficient when used alone in phylogenies, and can cause problems in phylogenetic studies (Sibley and Ahlquist, 1991), the high rates of evolution experienced by mitochondrial genes (Brown et al., 1979) when associated with other nuclear markers increases information that is relevant for species distinction (Clabaut et al., 2005). This fact, the increasing of information from the association of a nuclear and mitochondrial marker, may be observed for the *LWS* associated with *COI*.

The *SWS1* gene, which is also a member of the opsin family, has been associated with the phylogeny of large vertebrate groups. It is efficient for use to establish phylogenetic hypotheses. In fact, it has the advantage of presenting few indels and thus reducing ambiguity during alignment, and has a non-complex relationship of nucleotide replacement, showing a replacement process of homogenous bases that are associated with strong phylogenetic signs (van Hazel et al., 2006).

Consequently, because *LWS* is a member of the same gene family, it has common characteristics, with important polymorphisms in the two genes (Carleton, 2009; Carleton et al., 2010). *LWS* is the more variable (Terai et al., 2002) and is therefore of high use in distinctions up to the genus level.

LWS is also associated with several features related to behavior, such as sexual selection (van Hazel et al., 2006). It also has great ecological relevance (Terai et al., 2006) since the product it expresses (Spady et al., 2005) is under constant selection pressure (Terai et al., 2002; Smith et al., 2012). Consequently, the analysis of certain coding regions may be directly related to the selection mechanism and thus reflect the evolution history of the species (Schott et al., 2014).

The present results reveal that the *LWS* sequence is a highly efficient method to establish phylogenetic links of Neotropical cichlids studied at the genus level, although the efficiency is somewhat lower at the species level. However, when associated with *COI* sequences, these were efficient even in the establishment of phylogenies at the species level. Due to its strong relationship

with selection in visual fish, the use of this gene may be efficacious for the establishment of phylogenies of Neotropical cichlid species.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We would like to thank Programa de Ecologia de Ambientes Aquáticos Continentais (PEA), Núcleo de Pesquisa em Limnologia, Ictiologia e Aquicultura (Nupelia), and Universidade Estadual de Maringá (UEM) for the infrastructure made available for this research. We also appreciate Programa Ecológico de Longa Duração (PELD-CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the project funding and research scholarship. We are also grateful to Rodrigo Júnio da Graça for their contributions.

REFERENCES

- Amorim DS (2002). Fundamentos de Sistemática Filogenética. Holos, Riberirão Preto.
- Beheregaray LB (2008). Twenty years of phylogeography: The state of the field and the challenges for the Southern Hemisphere. *Mol. Ecol.* 17: 3754-3774.
- Brawand D, Wagner CE, Li YI, Malinsky M, et al. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature* 513: 375-381.
- Brown WM, George M and Wilson AC (1979). Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. U. S. A.* 76: 1967-1971.
- Buso GSC (2005). Marcadores Moleculares e Análise Filogenética. Embrapa, Brasília.
- Carleton K (2009). Cichlid fish visual systems: mechanisms of spectral tuning. *Integr. Zool.* 4: 75-86.
- Carleton KL, Hofmann CM, Klisz C, Patel Z, et al. (2010). Genetic basis of differential opsin gene expression in cichlid fishes. *J. Evol. Biol.* 23: 840-853.
- Clabaut C, Salzburger W and Meyer A (2005). Comparative phylogenetic analyses of the adaptive radiation of Lake Tanganyika cichlid fish: nuclear sequences are less homoplasious but also less informative than mitochondrial DNA. *J. Mol. Evol.* 61: 666-681.
- Darriba D, Taboada GL, Doallo R and Posada D (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9: 772.
- Drummond AJ, Suchard M a., Xie D and Rambaut A (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29: 1969-1973.
- Fabrin TMC, Simone I, Prioli SMAP, Prioli AJ, et al. (2014). A utilização de marcadores na filogenia dos ciclídeos (Teleostei: Perciformes): uma análise cienciométrica. *Enc. Biosf.* 10: 3118-3128.
- Gasques LS, Beloni KP and Oliveira JR (2013). Os marcadores moleculares em peixes e suas aplicações em publicações da base de dados Scielo. *Arq. Cienc. Vet. Zool. Unipar.* 16: 47-50.
- Gasques LS, Fabrin TMC, Prioli SMAP and Prioli AJ (2015). Prospecting of molecular markers for species of the *Cichla* genus. *Acta Sci. Biol. Sci.* 37: (In Press).
- Hillis DM (1987). Molecular versus morphological approaches to systematics. *Annu. Rev. Ecol. Syst.* 18: 23-42.
- López-Fernández H, Honeycutt RL and Winemiller KO (2005). Molecular phylogeny and evidence for an adaptive radiation of geophagine cichlids from South America (Perciformes: Labroidei). *Mol. Phylogenet. Evol.* 34: 227-244.
- López-Fernández H, Winemiller KO and Honeycutt RL (2010). Multilocus phylogeny and rapid radiations in Neotropical cichlid fishes (Perciformes: Cichlidae: Cichlinae). *Mol. Phylogenet. Evol.* 55: 1070-1086.
- Maan ME and Sefc KM (2013). Colour variation in cichlid fish: developmental mechanisms, selective pressures and evolutionary consequences. *Semin. Cell Dev. Biol.* 24: 516-528.
- Macrander JC (2010). Microsatellite Development, Population structure and demographic histories for two species of

- amazonian peacock bass *Cichla temensis* and *Cichla monoculus* (Perciforme: Cichlidae). Master's thesis. University of Nebraska, Lincoln.
- Miyagi R, Terai Y, Aibara M, Sugawara T, et al. (2012). Correlation between nuptial colors and visual sensitivities tuned by opsins leads to species richness in sympatric Lake Victoria cichlid fishes. *Mol. Biol. Evol.* 29: 3281-3296.
- Nagai H, Terai Y, Sugawara T, Imai H, et al. (2011). Reverse evolution in RH1 for adaptation of cichlids to water depth in Lake Tanganyika. *Mol. Biol. Evol.* 28: 1769-1776.
- Patwardhan A, Say S and Roy A (2014). Molecular Markers in Phylogenetic Studies-A Review. *J. Phylogenet. Evol. Biol.* 2: Doi: 10.4172/2329-9002.1000131.
- Ratnasingham S and Hebert PD (2013). A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. *PLoS One* 7: e66213.
- Rosenthal A, Coutelle O and Craxton M (1993). Large-scale production of DNA sequencing templates by microtitre format PCR. *Nucleic Acids Res.* 21: 173-174.
- Schott RK, Refvik SP, Hauser FE, López-Fernández H, et al. (2014). Divergent positive selection in rhodopsin from lake and riverine cichlid fishes. *Mol. Biol. Evol.* 31 :1149-1165.
- Schwanzara S (1967). The visual pigments of freshwater fishes. *Vision Res.* 7: 121-148.
- Sibley CG and Ahlquist JE (1991). Phylogeny and classification of the birds of the world. Yale University Press, New Haven.
- Silvestro D and Michalak I (2012). RaxmlGUI: A graphical front-end for RAxML. *Org. Divers. Evol.* 12: 335-337.
- Smith AR and Carleton KL (2010). Allelic variation in Malawi cichlid opsins: a tale of two genera. *J. Mol. Evol.* 70: 593-604.
- Smith AR, Ma K, Soares D and Carleton KL (2012). Relative LWS cone opsin expression determines optomotor thresholds in Malawi cichlid fish. *Genes Brain Behav.* 11: 185-92.
- Spady TC, Seehausen O, Loew ER, Jordan RC, et al. (2005). Adaptive molecular evolution in the opsin genes of rapidly speciating cichlid species. *Mol. Biol. Evol.* 22: 1412-1422.
- Steel M and Penny D (2000). Parsimony, Likelihood, and the Role of Models in Molecular Phylogenetics. *Mol. Biol. Evol.* 17: 839-850.
- Sullivan J, Swofford D and Naylor G (1999). The effect of taxon sampling on estimating rate heterogeneity parameters of maximum-likelihood models. *Mol. Biol. Evol.* 16: 1347-1356.
- Tamura K, Stecher G, Peterson D, Filipski A, et al. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30: 2725-2729.
- Terai Y, Mayer WE, Klein J, Tichy H, et al. (2002). The effect of selection on a long wavelength-sensitive (LWS) opsin gene of Lake Victoria cichlid fishes. *Proc. Natl. Acad. Sci. U. S. A.* 99: 15501-15506.
- Terai Y, Seehausen O, Sasaki T, Takahashi K, et al. (2006). Divergent selection on opsins drives incipient speciation in Lake Victoria cichlids. *PLoS Biol.* 4: e433.
- Thompson JD, Higgins DG and Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673-4680.
- Trezeise A and Collin S (2005). Opsins: evolution in waiting. *Curr. Biol.* 15: R794-R796.
- van Hazel I, Santini F, Müller J and Chang BSW (2006). Short-wavelength sensitive opsin (SWS1) as a new marker for vertebrate phylogenetics. *BMC Evol. Biol.* 6: 97.
- Weadick CJ, Loew ER, Rodd FH and Chang BSW (2012). Visual pigment molecular evolution in the Trinidadian pike cichlid (*Crenicichla frenata*): a less colorful world for neotropical cichlids? *Mol. Biol. Evol.* 29: 3045-3060.
- Wiley EO, Johnson GD and Dimmick WW (1998). The phylogenetic relationships of lampridiform fishes (Teleostei: acanthomorpha), based on a total-evidence analysis of morphological and molecular data. *Mol. Phylogenet. Evol.* 10: 417-425.