Artículo

Rev. Chilena Ent.
2015, 40: 14 - 21

THE PHYLOGENETIC POSITION OF *CEROGLOSSUS OCHSENII* GERMAIN AND *CEROGLOSSUS GUERINI* GERMAIN (COLEOPTERA: CARABIDAE), TWO ENDEMIC GROUND BEETLES FROM THE VALDIVIAN FOREST OF CHILE

POSICION FILOGENETICA DE *CEROGLOSSUS OCHSENII* GERMAIN Y *CEROGLOSSUS GUERINI* GERMAIN (COLEOPTERA: CARABIDAE), DOS ESCARABAJOS DE SUELO ENDEMICOS DEL BOSQUE VALDIVIANO DE CHILE

Carlos Muñoz-Ramírez [1]

ABSTRACT

The phylogenetic position of two *Ceroglossus* species, *C. ochsenii* Germain and *C. guerini* Germain, is evaluated for the first time based on molecular data. The results of this study showed that *C. ochsenii* and *C. guerini* are closely related to *C. suturalis*, with the three species clustering in a strongly supported clade in agreement with previous morphological studies. Although clearly differentiated from *C. suturalis*, these two species show low genetic divergence between them and do not form reciprocally monophyletic clades, rising questions on their status of separated biological species. Finally, and giving the recent proliferation of subspecific descriptions within *Ceroglossus*, I discuss the merits of these practices and call for caution on describing new species/subspecies to avoid the inflation of diversity within the genus.

Key words: mtDNA, Phylogeny, South Chile, Systematics.

RESUMEN

En este trabajo se evalúa por primera vez la posición filogenética de *Ceroglossus ochsenii* Germain y *Ceroglossus guerini* Germain dentro del género sobre la base de información molecular. Los árboles filogenéticos mostraron que estas dos especies se encuentran estrechamente emparentadas con *C. suturalis* en un clado con fuerte soporte nodal, en concordancia con trabajos morfológicos previos. Aunque claramente diferenciadas de *C. suturalis*, estas especies (*C. ochsenii* and *C. guerini*) presentaron baja divergencia genética y ausencia de monofilias reciprocas, levantando dudas sobre la validez de su designación como dos especies biológicas. Finalmente, y dada la proliferación de descripciones subespecíficas en la literatura reciente, se discute el mérito de estas designaciones y se llama a ser cautos a la hora de describir nuevas entidades para evitar la sobre-estimación de la diversidad biológica en *Ceroglossus*.

Palabras clave: Filogenia, mtDNA, Sistemática, Sur de Chile.

INTRODUCTION

The genus *Ceroglossus* Solier 1848 represents a remarkable group of flightless ground beetles endemic to the temperate forest of southern South America (Jiroux 2006). Widely distributed in forests of Central and South Chile, these carabids are highly variable intraspecifically, exhibiting different color morphs depending on the geographic area from where they are sampled. Due to this extraordinary

[1] Department of Ecology and Evolutionary Biology, Museum of Zoology, University of Michigan, Ann Arbor, MI 48109-1079, USA. E-mail: carmunoz@umich.edu

color diversity and yet relatively conserved general morphology, their taxonomic study has been particularly challenging. This is illustrated by the existence of more than one hundred names that have been used in the literature since the first description of a species in 1829 (Roig-Juñet & Domínguez 2001). The most recent revisions of the genus (Jiroux 1996, 2006), however, have reduced the number of species to only eight, although many of the previously available names have been maintained for many subspecific designations. Jiroux (1996) proposed a classification for *Ceroglossus* consisting of four main groups or lineages, based on morphological characters: *C. chilensis* (Eschsoltz 1829), *C. buqueti* (Laporte 1834), *C. darwini* group (i.e. containing *C. darwini* (Hope 1837), *C. magellanicus* Gehin 1885 and *C. speciosus* Gerstaecker 1858), and *C. suturalis* group (i.e. containing the species *C. suturalis* (Fabricius 1775), *C. ochsenii* (Germain 1895), and *C. guerini* (Germain 1895)). This systematic hypothesis, which relies mainly on characters from the aedeagus and the presence/absence and position of carenas in the antennae of males, was partially supported later by molecular phylogenetic analyses (Okamoto *et al*. 2001) which, by including six out of the eight known species, recovered the *C. darwini* group as monophyletic. The group "suturalis", however, was represented only by the species *C. suturalis* in the Okamoto *et al*. (2001) analysis, and the phylogenetic position of the two other species in the group, the rare *C. ochsenii* and *C. guerini*, remained untested. The study of *C. darwini* and *C. guerini* is important because these species are relatively rare and have small distribution ranges, which make them particularly vulnerable in terms of conservation (Jerez *et al*. 2015, Pizarro-Araya *et al*. 2012).

In this study, I use the mitochondrial DNA (mtDNA) marker COI to investigate the phylogenetic position of both *C. ochsenii* and *C. guerini* and to generally discuss and comment the findings in relation to previous work. Finally, I provide an opinion on the current tendency of describing new subspecies, and call for caution to avoid an overestimation of the taxonomic diversity within de genus.

MATERIAL AND METHODS

Sample collection and DNA sequencing

A total of 25 individuals from all eight species were analyzed in this study (Table 1). Three to six specimens were used per species except *C. ochsenii* (2) and *C. guerini* (1) for which only two and one specimens were available, respectively. For all species except *C. ochsenii* and *C. guerini*, individuals were chosen from geographically distant localities to better represent within-species diversity.

Genomic DNA was extracted from legs using the QIAGEN DNeasy Tissue Kit (QIAGEN Inc., Chatsworth CA) following the manufacturer protocol. For the only two dry specimens, one from *C. ochsenii* and one from *C. guerini*, the complete head and pronotum were used in the extraction procedure to ensure obtaining enough amplifiable DNA. A portion of the mtDNA COI gene (~680 bp), was amplified for all individuals using the universal primers LCO1490 5'- GGTCAACAAATCATAAAGATATTGG-3' and HCO 2198 5'-TAAACTTCAGGGTGACCAAAAAATCA-3' (Folmer *et al*. 1994). Each PCR reaction contained 1µl of extracted DNA, 2 µl of 10x buffer, 1.5µl of MgCl2, 1µl of 10mM dNTPs, 0.4 µl of 1% BSA, 0.8µl of each primer (10µM), 0.06µl of Tag DNA polymerase (Invitrogen, USA), and ddH2O to make a total of 25µl reaction. A standard PCR profile with one-minute duration for each step, a total of 35 cycles, and a final extension of 10 minutes at 72°C was followed. The annealing temperature was 52°C. PCR products were sequenced on an ABI Model 3730 XL sequencer by the Sequencing Core, University of Michigan, USA. All sequences were deposited in GenBank, with accession numbers KT997732, KT997737, KT997740, KT997744–45, KT997747, KT997750, KT997752–54, KT997760–

Table1: Locality data and voucher specimens used in this study.

Tabla 1: Localidades y especímenes usados en este estudio.

Locality	Latitude	Longitude	*C. chilensis*	*C. buqueti*	*C. darwini* group	*C. suturalis* group
1. Radal-Siete Tazas	-35.4756	-70.9938	ST.Cc.01			
2. Retiro	-36.0908	-71.7834	Ret.Cc.01			
3. Nahuelbuta	-37.8193	-73.0283	Nah1.Cc.01	Nah3.Cb.01	Nah3.Cm.01	
4. Malalcahuello	-38.471	-71.576			Mal5.Cm.01	
5. Pucón	-39.3508	-71.968	Puc.Cc.02	Puc.Cb.02	Puc.Cm.02	
6. Neltume	-39.8511	-71.9256			HH1.Cm.01	
7. Máfil	-39.7028	-72.9198				Maf.Cg.01
8. Parque Alerce Costero	-40.2063	-73.4129	Ale3.Cc.01		Ale2.Cd.01	Ale.Co.02; Ale4.Co.01
9. Puyehue	-40.6641	-72.172		Puy.Cb.01	Puy.Cd.01	
10. Ancud, Chiloé	-41.882	-73.8799			Chil1. Csp.01-03	
11. Puntra, Chiloé	-42.1195	-73.8066	Chil2.Cc.01			
12. Cucao, Chiloé	-42.6481	-74.0653				Chil5.Cs.01
13. Chaitén	-42.9097	-72.7074		Cb.Chai.01	Cd.Chai.01	
14. Aysén	-45.1327	-73.0154		AyRM. Cb.01		AyRM.Cs.01
15. Isla Navarino	-54.948	-67.646				Nav.Cs.01

61, KT997764–65, KT997768, and KT99777–82.

Sequence editing and Phylogenetic analyses

Chromatograms were edited in CodonCode Aligner version 3.0.3. and then imported in Bioedit 7.2.5 (Hall *et al*. 2011). Sequences were then aligned using the ClustalW algorithm implemented in Bioedit and checked via amino acid coding in MEGA 6.0 (Tamura *et al*. 2013) to test for unexpected frame shift errors or stop codons. Previous to the phylogenetic analyses, the Xia test (Xia *et al*. 2003) was conducted to evaluate the degree of sequence saturation, and therefore, the utility of the marker for phylogenetic reconstruction.

Phylogenetic relationships were estimated by maximum likelihood (ML) and Bayesian inference (BI) using the best-fit model selected by JModeltest (Posada 2008) under the Bayesian information criterion (BIC). JModeltest identified HKY+I+G as the best-fit model with the following parameter estimates: Lset Base = (0.3030 0.1709 0.1464) Nst = 2 Tratio = 5.8132 Rates = gamma Shape=1.6760 Ncat=4 Pinvar=0.6880. The ML tree was reconstructed using the software PAUP (Swofford 2002). The tree was estimated via a heuristic search with five random additions of taxa and TBR branch swapping. Node support was assessed by 1000 bootstrap replicates using a heuristic search with 5 random additions of taxa and TBR branch swapping. It is known that distant outgroups (relative to the ingroup) might not be suitable for rooting some phylogenetic trees (Graham *et al*. 2002). In this

case, *Ceroglossus* diverged from its closest relatives (e.g. the lineage containing *Carabus* and *Callosoma*) about 50 my ago (Andújar *et al*. 2012), which certainly makes the rooting by outgroup more challenging. Here, the tree was rooted via the midpoint method (after testing for the molecular clock assumption in PAUP) given that preliminary attempts of rooting via outgroup yielded problematic results (e.g. the root commonly fell among intraspecific clades).

The BI analysis was performed in MrBayes 3.1.2 (Ronquist & Huelsenbeck 2003) and posterior probabilities resulting from this analysis were added as branch support to the ML tree. Two independent runs were run to check for convergence of parameters. Four chains were used for phylogeny estimation, starting the analysis with a random tree and running it for 5,000,000 generations, sampling every 1000 trees and discarding the initial 20% as burning using the 'sumt' command. Convergence of parameters from the two independent runs was assessed by the average standard deviation of split frequencies (i.e. values below 0.01) and the potential scale reduction factor (PSRF; values should approach 1). Topological convergence was evaluated by using the plotting tool *compare* implemented in the online program AWTY (Nylander *et al*. 2008). Once convergence was confirmed, the two runs were combined to obtain a total of 8002 trees.

RESULTS

Xia's test showed no substitution saturation (Iss=0.3; Iss.c=0.73; p=0.000), which indicates the molecular data is useful for phylogenetic inference. Phylogenetic analyses showed four major clades in agreement with the four-species hypothesis proposed by Jiroux (1996) on the basis of morphological characters (ML best-tree score= 2607.362; Fig. 1, clades A-D): *C. chilensis* (clade A), *C. darwini* group (clade B), *C. buqueti* (clade C), and *C. suturalis* group (clade D). *Ceroglossus ochsenii* and *C. guerini* appeared more closely related with *C. suturalis* than to any other *Ceroglossus* species in agreement with the taxonomic scheme of Jiroux (1996). Support for all main clades was generally high, indicated either by high bootstrap values and/or high Bayesian posterior probabilities. A few other internal clades were also recovered with high support (subclades a1–a3, b1–b2, c1–c2 and d2). In addition, some internal subclades matched more or less some species boundaries. Subclade b1 corresponded to the species *C. magellanicus*,

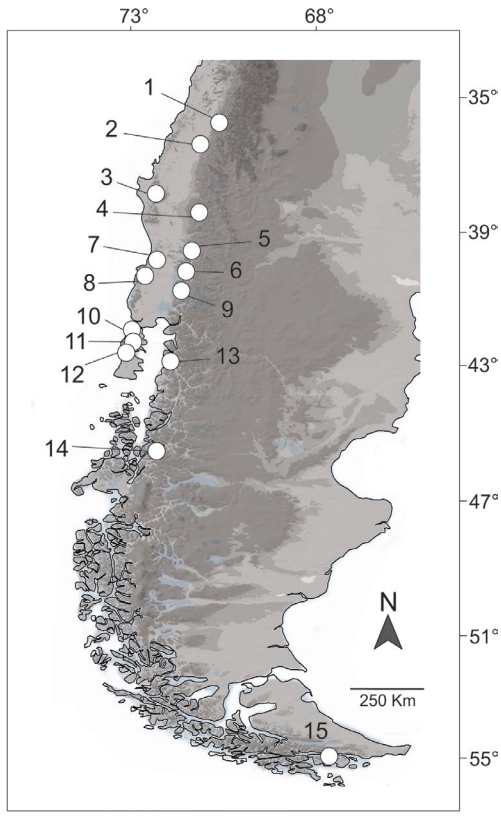

Figure 1: Sampling sites for *Ceroglossus* specimens used in this study. Information associated to these sites can be found in Table 1.

Figura 1: Localidades de colecta de los *Ceroglossus* analizados en este estudio. Datos asociados a estos puntos se encuentran en la Tabla 1.

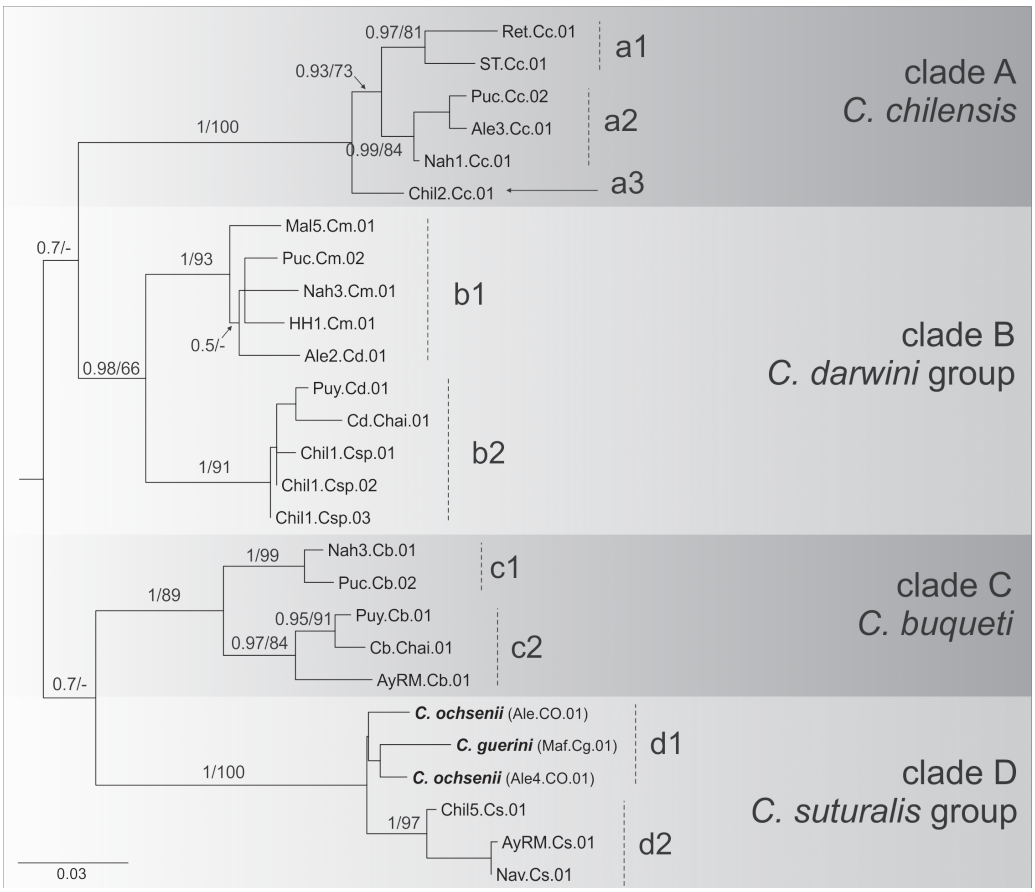

Figure 2: Maximum likelihood (likelihood score (-lnL) = 2613) phylogenetic reconstruction of *Ceroglossus* species. Branch support is indicated by the values above the branches (Bayesian posterior probability/ML bootstrap). Only values equal or above 0.5/50 are shown. See Table 1 for details about tip labels.

Figura 2: Árbol filogenético estimado para las especies de *Ceroglossus* mediante máxima verosimilitud (ML). El soporte de rama se indica sobre éstas (probabilidad a posteriori de Bayes/ML bootstrap). Sólo se muestran los valores mayores o iguales a 0.5/50. Ver Tabla 1 para mayores detalles sobre las etiquetas de las ramas terminales.

while clade d1 corresponded to the species *C. suturalis*. Subclade b2, on the other hand, contained both *C. darwini* and *C. speciosus*, with individuals from the latter being grouped within a weakly supported subclade. These species were not represented by reciprocally monophyletic clades, showing paraphyletic relationships. The individual Ale2.Cd.01 from Parque Alerce Costero (identified as *C. darwini ugartei* following Jiroux 2006) was more closely related to the species *C. magellanicus* than to the species *C. darwini*.

Within the clade D, that contains the *C. suturalis* group, the relationship between the three species was not resolved, and even though the species *C. suturalis* was recovered as a monophyletic group, *C. guerini* and *C. ochsenii* were not recovered as reciprocally monophyletic

groups. In other words, one individual from *C. ochsenii* (Ale4.Co.01; Table 1) appeared more closely related to *C. guerini* than to the other individual of *C. ochsenii* (Ale.CO.01).

DISCUSSION

The results support the taxonomic scheme of Jiroux (1996) that considered the existence of four major groups within the genus *Ceroglossus*. Please consider the relationships between the four major lineages with caution as the rooting was conducted by the midpoint rooting method because outgroup rooting resulted problematic due to the high evolutionary distance separating outgroup and ingroup (Andújat *et al.* 2012). However, relationships within each main lineages, which is the main purpose of this study, should not be affected by this issue. Here I focus on relationships within main lineages rather than between main lineages. As proposed by Jiroux (1996, 2006), *C. ochsenii* and *C. guerini* belong to the same group that contains *C. suturalis* to which they are sister species. However, *C. ochsenii* and *C. guerini* do not formed reciprocally monophyletic clades, indicating low genetic divergence between these species.

There are two possible explanations for the low divergence between *C. ochsenii* and *C. guerini*. One possibility is that these species have diverged very recently, so species still maintain some ancestral polymorphism (i.e. it has not been completely sorted by genetic drift). The other possibility is that these entities have not completed the speciation processes, and the lack of reciprocally monophyletic clades reflects ongoing or very recent gene flow. Morphological differences between *C. ochsenii* and *C. guerini* are well established in the literature (Jiroux 1996, 2006) and seem to support the former hypothesis. Following this argument, the lack of reciprocally monophyletic clades would be caused by the presence of incomplete lineage sorting (ILS), which can be common in recently diverged taxa that exhibit relatively large population sizes. However, these species are rare and have small geographic distributions, suggesting they have relatively small population sizes. Under small population sizes, the presence of ILS is less likely, as the higher magnitude of genetic drift in small population would accelerate the formation of reciprocally monophyletic clades. It is possible then, that the lack of monophyly observed in these species may reflect ongoing gene flow instead of ILS, and the observed morphological differences respond to phenotypic variation within species rather than difference between species. If that is the case, the validity of these entities as two separate biological species should be re-examined in the light of new evidence.

Low divergence was also found between *Ceroglossus darwini* and *Ceroglossus speciosus*. Previous phylogenetic work (Okamoto *et al*. 2001) included only one specimen of *C. speciosus*, so patterns of reciprocal monophyly (or the lack of it) could not be assessed. This species also have small geographic range (i.e. Ancud and surrounding areas) so questions remain whether the low divergence reflects recent speciation, recent gene flow, or both. A pattern that was consistent with previous molecular work is the placement of *C. darwini ugartei* into the clade of *C. magellanicus*. The inclusion of this subspecies within *C. darwini* was acknowledged as problematic by Jiroux (2006) due to the presence of characters from both species. The phylogenetic results from this study and from Okamoto *et al*. (2001) clearly show this subspecies belong to *C. magellanicus*, so amendments should be made to update the *Ceroglossus* taxonomy.

Comments on the phenotypic diversity within *Ceroglossus*

Although there is a high phenotypic diversity within some *Ceroglossus* species, this work does not provide a basis for distinguishing diversity at a finer taxonomic scale (i.e. subspecies or color morphs). Indeed, the lack of reciprocal monophyly found between some species suggests that making claims about the

validity of some species designations can be challenging even for species with clearly distinguishable phenotypes. Then, making the case for different taxonomic entities below the level of species should be done carefully, after examining different types of evidence (e.g. DNA and Morphology) (Dayrat 2005, Padial *et al*. 2010). Unfortunately, there has been a proliferation of work describing new subspecies in the recent literature, which do not rely on multiple sources of evidence and lack the use of clear biological criteria. This practice may distort the amount of diversity that actually exists within these taxa, and create problems related with the stability of its taxonomy. I am not against the practice of describing subspecific diversity, as long as it relies on biologically informed criteria and it is supported by multiple types of evidence (e.g. ecological, behavioral, morphological, geographical, or molecular). Describing new subspecies without the appropriate support can artificially inflate the actual diversity within this group and have a negative impact on studies that build on information about taxonomic diversity (e.g. ecological, biogeographical, and conservation studies). Further studies are needed to understand the mechanisms promoting phenotypic diversity in *Ceroglossus*, and future work should focus on testing the actual diversity within the genus. Until then, researchers should refrain from describing new species/subspecies without making reference to clearly defined criteria and analyzing multiple, independent types of evidence.

ACKNOWLEDGMENTS

I want to thank the Corporación Nacional Forestal (CONAF) for allowing sampling within several National Parks (Authorization number: 024/2012; see Table 1). I also appreciate the material donated by Álvaro Zúñiga, Dusan Boric, Gustavo Valenzuela, Mauricio Cid, Christian Muñoz, Laurent Rataj, and Dr. Viviane Jerez. I am especially grateful to Álvaro Zuñiga, Christian Muñoz, Alfonso Jara, Luis Correa, Jorge Avaria, Pablo Fuentes, and many others for their selfless assistance in the field and appreciate the comments and feedback provided by two anonymous reviewers that greatly improved the manuscript.

REFERENCES

Andújar, C., Serrano, J., & Gómez-Zurita, J. (2012) Winding up the molecular clock in the genus *Carabus* (Coleoptera: Carabidae): assessment of methodological decisions on rate and node age estimation. *BMC Evolutionary Biology, 12* (1), 40. doi:10.1186/1471-2148-12-40

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, *3* (5), 294 – 299. doi:10.1371/journal.pone.0013102

Graham, S. W., Olmstead, R. G., & Barrett, S. C. H. (2002) Rooting phylogenetic trees with distant outgroups: a case study from the commelinoid monocots. *Molecular Biology and Evolution*, *19* (10), 1769 – 1781. doi:10.1093/oxfordjournals.molbev.a003999

Hall, T., Biosciences, I., & Carlsbad, C. (2011) BioEdit: An important software for molecular biology. *GERF Bulletin of Biosciences*, *2*, 60 – 61.

Jerez, V., Zúñiga-Reinoso, Á., Muñoz-Escobar, C., & Pizarro-Araya, J. (2015) Acciones y avances sobre la conservación de insectos en Chile. *Gayana* (Concepción), *79* (1), 1 – 3. doi:10.4067/S0717-65382015000100001

Jiroux, E. (1996) Revision du genre Ceroglossus, Vol. 1. *Collection Systematique*. Verneuil-sur-Seine: Magellanes.

Jiroux, E. (2006) Le genre Ceroglossus, Vol. 14. *Collection Systematique*. Verneuil-sur-Seine: Magellanes.

Nylander, J. A. A., Wilgenbusch, J. C., Warren, D. L., & Swofford, D. L. (2008) AWTY (are we there yet?): A system for graphical exploration of MCMC convergence in Bayesian

phylogenetics. *Bioinformatics*, *24* (4), 581 – 583. doi:10.1093/bioinformatics/btm388

Okamoto, M., Kashiwai, N., Su, Z. H., & Osawa, S. (2001) Sympatric convergence of the color pattern in the Chilean *Ceroglossus* ground beetles inferred from sequence comparisons of the mitochondrial ND5 gene. *Journal of Molecular Evolution*, *53* (4 - 5), 530–8. doi:10.1007/s002390010243

Pizarro-Araya, J., Vergara, O. E., & Flores, G. E. (2012) *Gyriosomus granulipennis* Pizarro-Araya & Flores 2004 (Coleoptera: Tenebrionidae): Un caso extremo a conservar. *Revista Chilena de Historia Natural, 85* (3), 345 – 349. doi:10.4067/S0716-078X2012000300009

Posada, D. (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*. doi:10.1093/molbev/msn083

Roig-Juñet, S., & Dominguez, M. C. (2001) Diversidad de la familia Carabidae (Coleoptera) en Chile. *Revista Chilena de Historia Natural, 74*, 549 – 571.

Ronquist, F., & Huelsenbeck, J. P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19, 1572 – 1574. doi:10.1093/bioinformatics/btg180

Swofford, D. L. (2002) PAUP* phylogenetic analysis using parsimony (*and other methods). Version 4.0b10. *Sinauer* Associates.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), 2725–2729. doi:10.1093/molbev/mst197

Xia, X., Xie, Z., Salemi, M., Chen, L., & Wang, Y. (2003) An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution, 26* (1), 1 – 7. doi:10.1016/S1055-7903(02)00326-3.

(Recibido: 10 agosto 2015; Aceptado 3 diciembre 2015; Publicado: 30 diciembre 2015)
Manejado por Viviane Jerez.