

Genetic identification of benthic polychaetes in a biodiversity hotspot in the southeast Pacific

Identificación genética de poliquetos bentónicos en un hotspot de biodiversidad del Pacífico suroriental

Cristian B. Canales-Aguirre^{1,2}, Nicolás Rozbaczylo³
and Cristián E. Hernández¹

¹Laboratorio de Diversidad Molecular y Filoinformática, Departamento de Zoología, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Casilla 160-C, Concepción, Chile. cristiancanales@udec.cl

²Laboratorio de Genética y Acuicultura, Departamento de Oceanografía, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Casilla 160-C, Concepción, Chile

³Departamento de Ecología, Facultad de Ciencias Biológicas, Pontificia Universidad Católica de Chile, Santiago, Chile

Abstract. - The aim of this study was to evaluate the utility of COI, of the mitochondrial DNA, for the taxonomic determination of soft-bottom benthic polychaetes in the Chiloé Inner Sea as a particularly important zone for polychaete fauna of the southeast Pacific. The results show high genetic differentiation between morphological species of polychaetes (*posterior probabilities* upper 0.9), and *Onuphis pseudoiridescens* showed the greatest divergence (3.08 ± 0.08 S.D.). The results support the role of the COI gene as a useful molecular marker for fast and accurate taxonomic determination of benthic polychaetes. This study leads the way for research on the biodiversity and systematics of polychaetes off the Chilean coast using molecular phylogenetic criteria.

Key words: Genetic barcode, biodiversity, molecular taxonomy, phylogenetic systematic

INTRODUCTION

DNA genetic barcode can be of significant help for taxonomical, ecological and biological studies, mainly for specific identification in research on biological communities and biodiversity in general (Valentini *et al.* 2009). The main objective of using the DNA barcode is the creation of reference databases with nucleotide sequences of currently known species (Ratnasingham & Hebert 2007). At the present time, the utility of COI has been verified in different vertebrates (*e.g.*, Ward *et al.* 2005) and invertebrate taxa (*e.g.*, Costa *et al.* 2007). These studies have significantly contributed to the characterization of biodiversity, revealing new species and resolving cryptic or morphologically similar species.

Given the high diversity and frequently complex taxonomic characterization of the class Polychaeta and, beyond that the Polychaeta are a paraphyletic group, the application of a genetic barcode could contribute to the description of biodiversity patterns and the evaluation of the phylogenetic processes with which the polychaetes diversified. The most widely used molecular markers for polychaetes in relationship evaluations at both an intraspecific and interspecific level, are the 18S and the 28S genes from nuclear DNA (Bleidorn *et al.* 2003), and the

COI and the 16S genes from mitochondrial DNA (*e.g.*, Jolly *et al.* 2006, Iannotta *et al.* 2007). Nevertheless, there are currently wide biogeographic zones such as the southeast Pacific, in which polychaete identifications have mainly been based on morphological traits (*e.g.*, Rozbaczylo *et al.* 2005, 2006a, b), and in which there are no studies of molecular taxonomy, phylogenetic relationships or phylogeography to date. This is an alarming fact, if we consider that 556 benthic and 21 pelagic polychaetes species have been identified from the Chilean southeast Pacific coast alone (Rozbaczylo & Moreno 2006), making the use of molecular markers essential for biogeographical and evolutionary characterization of this biodiversity. Latitudinally, the biodiversity of polychaetes off the southeast Pacific coast is characterized by two biogeographical provinces (*i.e.*, Peruvian and Magellanic) (Hernández *et al.* 2005), the Chiloé Archipelago being the geographical distribution limit of various species and an ecotonal zone between the two provinces (Hernández *et al.* 2005, Moreno *et al.* 2006). Additionally, this zone is considered a biodiversity hotspot, in which the richness of endemic species of polychaetes is exceptionally high (Moreno *et al.* 2006). Therefore, the main objective of the present study was to evaluate the utility of

COI, of the mitochondrial DNA, for the taxonomic determination of soft-bottom benthic polychaetes in the Chiloé Inner Sea as a particularly important zone for polychaete fauna of the southeast Pacific.

MATERIALS AND METHODS

The polychaetes were obtained in July 2006 as part of benthic samples taken on the CIMAR-12 Fiordos cruise aboard the Chilean navy oceanographic vessel, AGOR 'Vidal Gormaz' in the Chiloé Inner Sea (southern Chile). The stations that were sampled are: Golfo de Ancud (41°58.70S, 72°58.55W); Islas Desertores (42°19.90S, 72°57.12W); Seno de Reloncaví (41°40.43S, 72°47.68W); Islas Desertores (42°53.85S, 72°53.46W); Estuario de Reloncaví (41°42.88S, 72°38.48W); Estuario de Reloncaví (41°39.68S, 72°22.45W); Estuario de Reloncaví (41°34.66S, 72°19.97W); Estuario de Reloncaví (41°31.09S, 72°18.93W); Castro (42°31.09S, 73°46.81W); y Dalcahue - Quinchao (42°22.69S, 73°31.71W), with depths ranging from 43 to 458 meters (Stations CL-3 and 4, respectively). The samples were obtained with a box corer of 0.025 m² and preserved in absolute alcohol for subsequent molecular taxonomy studies.

The taxonomic determination of the species was based on the work of Hartmann-Schröder (1965), Montiel *et al.* (2004), and Rozbaczylo *et al.* (2005, 2006a, b). The taxa were grouped into clades according to Rouse & Fauchald (1997) and Rouse & Pleijel (2001). The species included for molecular identification were: *Cistenides ehlersi* (Hessle, 1917), *Leanira quatrefagesi* (Kinberg, 1855), *Maldane chilensis* (Hartmann-Schröder, 1965), *Maldane* sp., *Ninoleptognatha* (Ehlers, 1900), *Onuphis pseudoiridescens* (Averincev, 1972) and *Melinna cristata australis* (Sars 1851 as characterized by Hartmann-Schröder 1965).

The extraction of genomic DNA was carried out using standard CTAB 2X protocol by means of digestion with proteinase K, extraction with chloroform and precipitation with absolute ethanol. The DNA was re-suspended in ultrapure water (GIBCO Technology™). The universal primers LCO1490 and HCO2198 from Folmer *et al.* (1994) were used for the amplifications. The COI fragment (~650 pb) was amplified to a final volume of 30 µL, which contains 1 µL of template DNA, 3 mM MgCl₂, 0.125 mM dNTPs, 0.1 µM of each primer, and 1.25 U/µL of Taq DNA polymerase (Fermentas®) with the KCl buffer (*i.e.*, 100 mM Tris-HCl pH 8.8 at 25°C, plus 500 mM KCl, 0.8% v/v Nonidet P40). The amplification was carried out in a PCT-200 thermal cycler (MJ-Research®) with the following cycling conditions: an initial denaturation at 94°C for 1 min,

followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 49°C for 55 s, extension at 72°C for 90 s and a final extension at 72°C for 10 min. The samples were sequenced at the MacroGen® Company (Korea) in an automated DNA sequencer (Model 3730xl; Applied Biosystems). The sequences have been deposited in the genbank database under the accession number GQ229112-GQ229122.

The COI sequences obtained from the specimens were edited and aligned visually and using the SEQUENCHER v4.8 (GeneCodes Corp.) program. We performed a test introduced by Xia *et al.* (2003) to measure substitution saturation in a set of aligned nucleotide sequences, to evaluate whether these sequences are useful for phylogenetic analyses. The sequence alignments were used to generate a genetic distance matrix by sequence pairs using Maximum Composite Likelihood method (Tamura *et al.* 2004) based on Tamura-Nei model (TN93) (Tamura & Nei 1993) implemented in MEGA v4.0 software (Tamura *et al.* 2007). Markov chain Monte Carlo was used within a Bayesian framework (BMCMC) for phylogenetic reconstruction of different taxa following the methodology used by Rodríguez-Serrano *et al.* (2008). The phylogenetic analysis described was conducted first using only the sequences obtained in the present study plus an out-group of a related taxon (*i.e.*, *Sipunculus nudus* (Linnaeus, 1766), Class Sipunculidea). Thereafter our sequences were compared to a data base of 106 selected species from GenBank. The sequence selection criteria were: (1) that these sequences were similar in length (size in pb) to that obtained in the present study, thus optimizing the number of homologue characters used; and (2) included the greatest number possible of representative members of each morphological clade, so as to optimize the species allocation and strengthen the phylogenetic analysis results.

RESULTS AND DISCUSSION

The sequence size had a variation of ~675-681pb. The divergence between the sequence-pairs used in this study is shown in Table 1. The species *O. pseudoiridescens* presented the greatest divergence between sequence pairs with an average divergence value of 3.08 (± 0.08 S.D.) substitutions per site, whereas the comparisons between the rests of the pairs of species showed an average divergence of 0.70 (± 0.28 S.D.) substitutes per site. The substitution saturation test demonstrated that our sequences have little saturation (Iss = 0.524 < Iss.c = 0.741; *P* < 0.0001), thus validating their use for phylogenetic inference.

Table 1. Evolutionary divergence estimation between sequence pairs (COI) of polychaetes by Maximum Composite Likelihood method and TN93 model / Estimación de divergencia evolutiva entre pares de secuencias (COI) de poliquetos mediante el método de Máxima Verosimilitud Compuesta y modelo TN93

	1	2	3	4	5	6	7	8	9	10	11
1											
2	0.022										
3	0.918	0.903									
4	0.919	0.904	0.011								
5	0.902	0.887	0.019	0.025							
6	0.789	0.796	0.909	0.914	0.909						
7	0.918	0.927	0.762	0.763	0.777	0.747					
8	0.837	0.845	0.753	0.754	0.759	0.719	0.067				
9	0.728	0.756	0.848	0.833	0.854	0.653	0.708	0.678			
10	0.734	0.750	0.842	0.827	0.848	0.648	0.702	0.673	0.003		
11	3.155	3.102	3.041	3.062	3.013	3.061	3.226	3.157	3.003	2.997	

1: *Maldane* sp.; 2: *Maldane chilensis*; 3-5: *Leanira quatrefagesi*; 6: *Melinna cristata*; 7-8: *Cistenides ehlersi*; 9-10: *Ninoe leptognatha*; 11: *Onuphis pseudoiridescens*

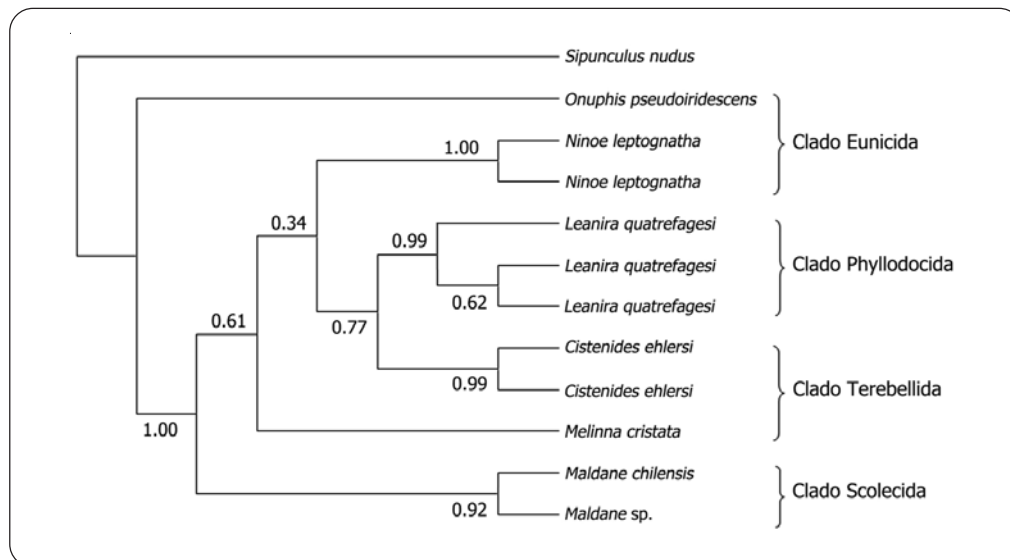


Figure 1. The consensus tree of the 4977 phylogenetic trees obtained in the convergence zone of the Markov chain by means of the Bayesian approach. The values above the nodes correspond to the posterior probability of the nodes / Árbol consenso de los 4977 árboles filogenéticos obtenidos en la zona de convergencia de las cadenas de Markov mediante aproximación bayesiana. Los valores sobre los nodos corresponden a la probabilidad *a posteriori* de los nodos

The phylogenetic tree constructed using the Bayesian approach, shows that the analyzed species present high values of posterior probability and they can be assigned to a defined clade (Fig. 1): *M. chilensis* and *Maldane* sp. (clade Scolecida); *L. quatrefagesi* (clade Phyllodocida); *N. leptognatha* and *O. pseudoiridescens* (clade Eunicida); and *C. ehlersi* and *M. cristata australis* (clade Terebellida). However, these last two clades do not group themselves into a monophyletic group. When evaluating the molecular taxonomic position of the species sequenced in this study

with those from the GenBank data base (Fig. 2), the samples sequenced in this study continued to show a high posterior probability and group themselves into species of a specific clade, but not necessarily into species of the same large morphological clade.

The present study is the first taxonomic identification based on the genetic barcode using the COI gene of polychaeta from a hotspot area in the southeast Pacific. The results from the genetic characterization show high values

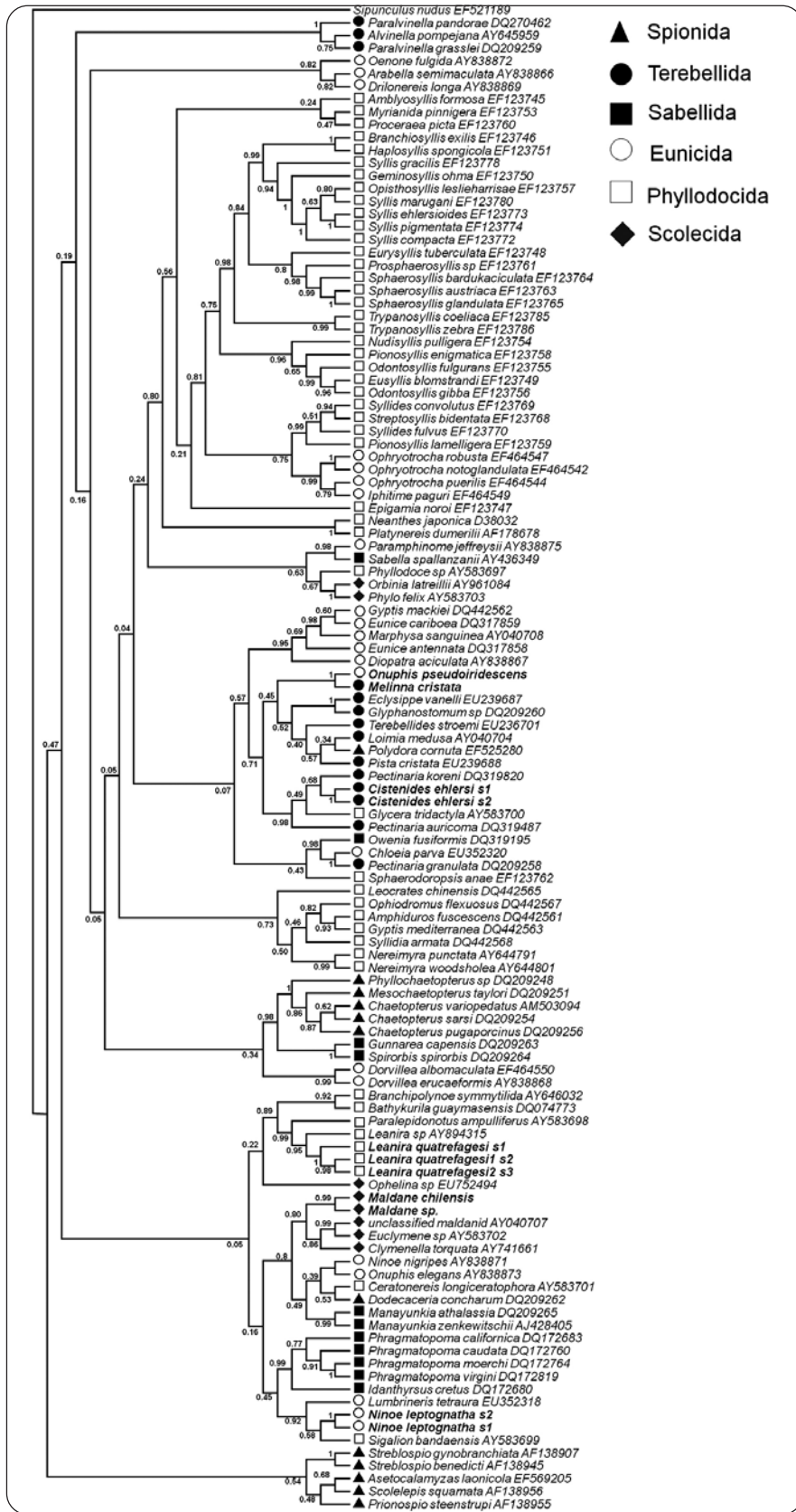


Figure 2. The consensus tree from the Bayesian approach used to compare with the data base obtained from Genbank. The values above the nodes correspond to the posterior probability. To the right of the species name is the access code that identifies the sequence in Genbank. Species in bold sequenced in this study / Árbol consenso obtenido de la aproximación bayesiana y usada para comparar con la base de datos obtenida desde Genbank. Los valores sobre los nodos corresponden a la probabilidad *a posteriori*. A la derecha de cada nombre de la especie se encuentra el código de acceso que identifica a la secuencia en Genbank. Especies en negrita fueron secuenciadas en este estudio

of divergence (Table 1) and posterior probability that support the morphology-based taxonomic delimitation (Fig. 1). These results are not unusual given that the species are very different in morphology, but in advance we didn't have any evidence about the genetic differentiation that we find in this study. On the other hand, the large morphological clades to which the species were assigned do not appear to be monophyletic (Fig. 1), which confirm the polyphyletic origin of morphological clades in polychaetes showed in two previous wide phylogenetic analysis performed by Hall *et al.* (2004) using the 18S ribosomal DNA gene sequences, and Rousset *et al.* (2007) using the 18S and the 28S ribosomal DNA D1 region sequences. For example, although various authors have carried out molecular phylogenetics reconstructions of polychaeta with the ribosomal gene 18S (Bleidorn *et al.* 2003, Hall *et al.* 2004, Rousset *et al.* 2007, Lehrke *et al.* 2007), recent studies show that this gene is saturated at the annelids level (Struck *et al.* 2008) and consequently it would not be useful for phylogenetics research at all taxonomic levels (Xia *et al.* 2003). However, our dataset has a little saturation.

Despite this, and given the consistency of the nodes that sustain the species in this study, the results corroborate the usefulness of the COI gene as a molecular marker for fast and accurate taxonomic determination of benthic polychaetes in the southeast Pacific, nevertheless at the taxonomic higher levels it tends to fall apart. This comparison contributes to the knowledge of polychaeta biodiversity by establishing a molecular data base that will help in the development of a well-founded taxonomic identification process and the future description of the evolutionary history of this group based on molecular phylogenetic approaches and evolution models based on the data.

Species identification in particular, in addition to phylogenetic relationship characterization in polychaeta, is a complex task which requires the combined work of classical and molecular taxonomists, especially because the class Polychaeta is the group with greatest species richness among the annelids (Rouse & Pleijel 2001, Rozbaczylo & Moreno 2006). In fact, in terms of diversity and abundance, the class Polychaeta is one of the most important taxa in soft-bottom benthic marine communities in the world (Knox 1977, Hutchings 1998). There are 9,000 nominal polychaete species described worldwide, grouped into 12 large morphological clades (Rouse & Pleijel 2001). There are 556 species off the Chilean coast alone described using mainly morphological traits (Rozbaczylo *et al.* 2005, Rozbaczylo *et al.* 2006a, b). Despite the great diversity of morphologically described polychaeta off the Chilean coast,

ecological, biogeographical and evolutionary knowledge is still incomplete (Rozbaczylo & Moreno 2006). Thus, the contribution of new data on molecular markers such as the COI will encourage the execution of further studies to help discover new patterns that underlie this great diversity and the processes that produce and maintain it.

Finally, the use of a DNA barcode for taxonomic identification can clearly help improve biodiversity classification and allow the critical examination of morphological traits commonly used in taxonomy (Valentini *et al.* 2009). The sampling conducted on the CIMAR-12 cruise in the Tenth Region of Chile, from the Reloncaví Sound (Puerto Montt) to Boca del Guafo (South of Chiloé), has allowed us to continue to acquire knowledge on marine biodiversity from a traditional morphological point of view, and begin a molecular systematic approach to benthic subtidal polychaeta in the biodiversity hotspot zone of this group off the southeast Pacific coast (Hernández *et al.* 2005, Moreno *et al.* 2006).

ACKNOWLEDGMENTS

The authors thank the Comité Oceanográfico Nacional (CONA) and the AGOR «Vidal Gormaz» crew, for their logistical support, especially Paz Pantoja, chief scientific officer during the first stage of the cruise. They also thank Cristóbal Bravo M., from Universidad de Concepción, for obtaining the samples on board and Patricio Zavala, from the Ecology Department, Pontificia Universidad Católica de Chile, for the processing of the samples. They also thank Sandra Sampertegui and Daniela Sepúlveda for their significant contributions to the manuscripts. This work is part of the CONA-C12F 06-06 Project, «Specific Diversity and Genetic Characterization Study of Benthic Polychaetes between Puerto Montt and Boca del Guafo (41°50's-43°45's), Chile» funded by the Servicio Hidrográfico y Oceanográfico de la Armada de Chile (SHOA). C.B. Canales-Aguirre is a student in the Doctoral Program in Systematics and Biodiversity, and was supported by UdeC Doctoral Fellowships. C.E Hernández also wishes to express his thanks for the partial funding from the DIUC - 207.113.098-1.0 Project, and FONDECYT-11080110 for the use of the computational platform.

LITERATURE CITED

Bleidorn C, L Vogt & T Bartolomaeus. 2003. New insights into polychaete phylogeny (Annelida) inferred from 18S rDNA sequences. *Molecular Phylogenetics and Evolution* 29: 279-288.

- Costa FO, JR deWaard, J Boutilier, S Ratnasingham, RT Dooh, M Hajibabaei & PD Hebert. 2007.** Biological identifications through DNA barcodes: the case of the Crustacea. *Canadian Journal of Fisheries and Aquatic Sciences* 24: 272-295.
- Folmer O, M Black, W Hoeh, R Lutz & R Vrijenhoek. 1994.** DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* 3: 294-299.
- Hall KA, PA Hutchings & DJ Colgan. 2004.** Further phylogenetic studies of the Polychaeta using 18S rDNA sequence data. *Journal of the Marine Biological Association of the United Kingdom* 84: 949-960.
- Hartmann-Schröder G. 1965.** Zur Kenntnis des Sublitorals der chilenischen Küste unter besonderer Berücksichtigung der Polychaeten und Ostracoden. Tl. II. Die Polychaeten des Sublitorals. *Mitteilungen aus dem Hamburgischen Zoologischen Museum und Institut* 62: 59-305.
- Hernández CE, RA Moreno & N Rozbaczylo. 2005.** Biogeographical patterns and Rapoport's rule in southeastern Pacific benthic polychaetes of the Chilean coast. *Ecography* 28: 363-373.
- Hutchings P. 1998.** Biodiversity and functioning of polychaetes in benthic sediments. *Biodiversity and Conservation* 7: 1133-1145.
- Iannotta MA, FP Patti, M Ambrosino, G Procaccini & MC Gambi. 2007.** Phylogeography of two species of Lysidice (Polychaeta, Eunicidae) associated to the seagrass *Posidonia oceanica* in the Mediterranean Sea. *Marine Biology* 150: 1115-1126.
- Jolly MT, F Viard, F Gentil, E Thiébaud & D Jollivet. 2006.** Comparative phylogeography of two coastal polychaete tubeworms in the Northeast Atlantic supports shared history and vicariant events. *Molecular Ecology* 15: 1841-1855.
- Knox GA. 1977.** The role of polychaetes in benthic soft-bottom communities. In: Reish DJ & K Fauchald (eds). *Essays on polychaetous annelids in memory of Dr. Olga Hartman*. pp. 547-604 Allan Hancock Foundation, Los Angeles.
- Lehrke J, HA ten Hove, TA Macdonald, T Bartolomaeus & C Bleidorn. 2007.** Phylogenetic relationships of Serpulidae (Annelida: Polychaeta) based on 18S rDNA sequence data, and implications for opercular evolution. *Organisms Diversity & Evolution* 7: 195-206.
- Moreno RA, CE Hernández, MM Rivadeneira, MA Vidal & N Rozbaczylo. 2006.** Patterns of endemism in south-eastern Pacific benthic polychaetes of the Chilean coast. *Journal of Biogeography* 33: 750-759.
- Rodríguez-Serrano E, RE Palma & CE Hernández. 2008.** The evolution of ecomorphological traits within the Abrothrichini (Rodentia: Sigmodontinae): A Bayesian phylogenetic approach. *Molecular Phylogenetics and Evolution* 48: 473-480.
- Rouse GW & K Fauchald. 1997.** Cladistics and polychaetes. *Zoologica Scripta* 26: 139-204.
- Rouse GW & F Pleijel. 2001.** *Polychaetes*, 354 pp. Oxford University Press, Oxford.
- Rousset V, F Pleijel, GW Rouse, C Erséus & ME Siddall. 2007.** A molecular phylogeny of annelids. *Cladistics* 23: 41-63.
- Rozbaczylo N & RA Moreno. 2006.** Polychaeta. In: Saball P, MTK Arroyo, JC Castilla, C Estades, JM Ladrón-Guevara, S Larraín, CA Moreno, F Rivas, J Rovira & L Sierralta (eds). *Biodiversidad de Chile. Patrimonio y desafíos*, pp. 175-178. Comisión Nacional del Medio Ambiente, Santiago.
- Rozbaczylo N, RA Moreno & O Díaz-Díaz. 2005.** Poliquetos bentónicos submareales de fondos blandos de la región de Aysén, Chile: Clado Phyllodocida (Annelida, Polychaeta). *Investigaciones Marinas* 33: 69-89.
- Rozbaczylo N, RA Moreno & O Díaz-Díaz. 2006a.** Poliquetos bentónicos submareales de fondos blandos de la región de Aysén, Chile: Clados Amphinomida, Eunicida, Spionida, Sabellida y Scolecida (Annelida, Polychaeta). *Investigaciones Marinas* 34: 43-62.
- Rozbaczylo N, RA Moreno, O Díaz-Díaz & S Martínez. 2006b.** Poliquetos bentónicos submareales de fondos blandos de la región de Aysén, Chile: Clado Terebellida (Annelida, Polychaeta). *Ciencia y Tecnología del Mar* 29: 71-90.
- Struck TH, MP Nesidal, G Purschke & KM Halanych. 2008.** Detecting possibly saturated positions in 18S and 28S sequences and their influence on phylogenetic reconstruction of Annelida (Lophotrochozoa). *Molecular Phylogenetics and Evolution* 48: 628-645.
- Tamura K & M Nei. 1993.** Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* 10: 512-526.
- Tamura K, M Nei & S Kumar. 2004.** Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences of the United States of America* 101: 11030-11035.
- Tamura K, J Dudley, M Nei & S Kumar. 2007.** MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 24: 1596-1599.
- Valentini A, F Pompanon & P Taberlet. 2009.** DNA barcoding for ecologists. *Trends in Ecology and Evolution* 24(2): 110-117.
- Ward RD, TS Zemlak, BH Innes, PR Last & PDN Hebert. 2005.** DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B, Biological Sciences* 360: 1847-1857.
- Xia X, Z Xie, M Salemi, L Chen & Y Wang. 2003.** An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution* 26: 1-7.

Received 14 March 2010 and accepted 23 November de 2010