Contents lists available at ScienceDirect

Data in Brief

Data Article

Nucleotide sequence and results of test of adaptive evolution in the α-globin gene of octodontoid rodents

I.H. Tomasco [a,*], N. Boullosa [a], F.G. Hoffmann [b,c], E.P. Lessa [a]

[a] Departamento de Ecología y Evolución, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo 11400, Uruguay
[b] Department of Biochemistry and Molecular Biology, Mississippi State University, MS, USA
[c] Institute for Genomics, Biocomputing and Biotechnology, Mississippi State University, MS, USA

ARTICLE INFO

ABSTRACT

The data presented in this article are related to the research article entitled "Molecular adaptive convergence in the α-globin gene in subterranean octodontid rodents" (Tomasco et al., 2017) [1]. This article shows the nucleotide sequences of α-globin subunit gene of hemoglobin of several South American caviomorph rodents, including subterranean and fossorial species. These sequences are deposited in Genbank, with accession numbers ranging from MF169881 to MF169898. Of a total of 429 nucleotides analyzed (143 codons), 100 variable sites and 43 amino acid replacements were reported. In this article we also show the results of TreeSaap (Woolley et al., 2003) [2] and MEME (Murrell et al., 2012) [3], that identified some replacement changes as interesting for future studies of adaptive evolution in this large rodent radiation.

Specifications Table

Subject area	*Biology*
More specific subject area	*Molecular Evolution*
Type of data	*Figures and tables*
How data was acquired	*Standard sequencing, and MEME (datamonkey.org), TreeSAAP software*
Data format	*Analyzed*
Experimental factors	*Edited sequences and use of Datamonkey server and TreeSaap Softwares*
Experimental features	*Standard sequencing from PCR products, edition and analyses.*
Data source location	*South America*
Data accessibility	*Genbank: accession from MF169881 to MF169898*

Value of the data

- Adaptations to underground life have begun to be studied in some groups of mammals. In this article we show and analyze the nucleotide variation of endemic South American rodents, where two relatively recent occupation of subrerranean life allows an analyses within a robust phylogenetic framework, comparing fully subterranean species with their closest fossorial relatives.
- In these data, 43 amino acid replacements of the α-globin gene of 7 species of caviomorph rodents are reported for the first time, belonging to the family Ctenomyidae (Ctenomys rionegrensis, C. leucodon, C. sociabilis), Octodontidae (Spalacopus cyanus, Aconaemys fuscus, Tympanoctomys barrerae), Echimyidae (Proechimys longicaudatus) and Caviidae (Cavia aperea), which is a significant contribution to the diversity of these rodents, hitherto represented by three species (i.e.: Cavia porcellus, Chinchilla lanigera and Octodon degus).
- In a related a research article [1] we showed at least two sites under positive selection in the basal branch of Octodontidae. However, further studies are needed to identify the real role of those amino-acid replacement selected, such as biochemical and physiological assays activity in different conditions.
- We also show results of TreeSAAP and MEME software that could, that together with biochemical and physiological studies and/or ideally complemented with site-directed mutagenesis, to shed light on the effect of those changes. So, this results may be a starting point for such research, and generate new biological hypotheses for mutation studies and functional analyses.
- Finally, sampling the large number of tuco-tuco species would provide more statistical power and increase the ability to examine adaptive changes along multiple lineages.

1. Data

A brief and conservative analysis of these results has been presented in a research article submitted to Gene journal [1]. In that study, the authors showed that at least sites 57 and 71 of the alpha globin gene show evidence of having been positively selected in the evolution of the octodontoid rodents. The door is left open for the analysis of other sites, for which the evidence was not so robust and are presented in this article. Particularly, summarized the complete cds sequences (Fig. 1) and all significant results obtained with MEME (Fig. 2) and TreeSAAP (Table 1) software. Particularly, MEME results show an interesting non-significant results is the site 111, that have nonsynonymous changes in the basal branch of subterranean Ctenomys (Fig. 2), and. TreeSAAP results show very radical changes in several branches of the tree (Table 1). The biological meaning of the changes in all these codon sites should be analyzed in the future in more details.
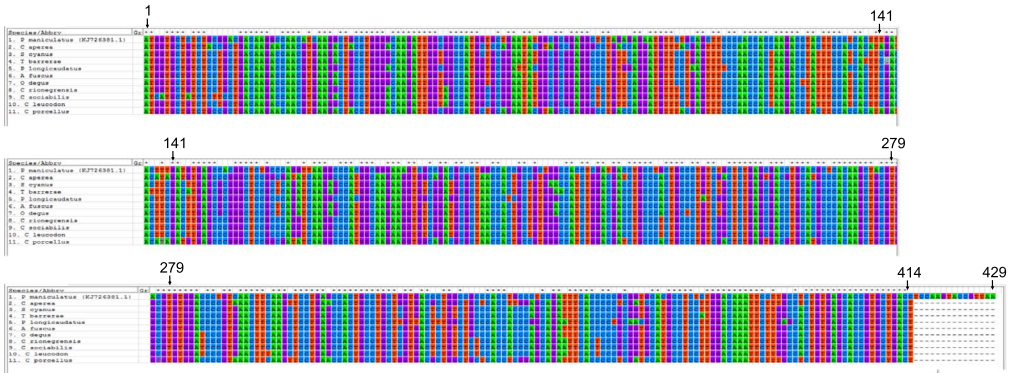
Fig. 1. CDS of alpha globin gene of hemoglobin. These sequences were obtained in this study for several South American rodents. The sequence of *Peromyscus maniculatus* was added as reference. The position in the alignment was included.
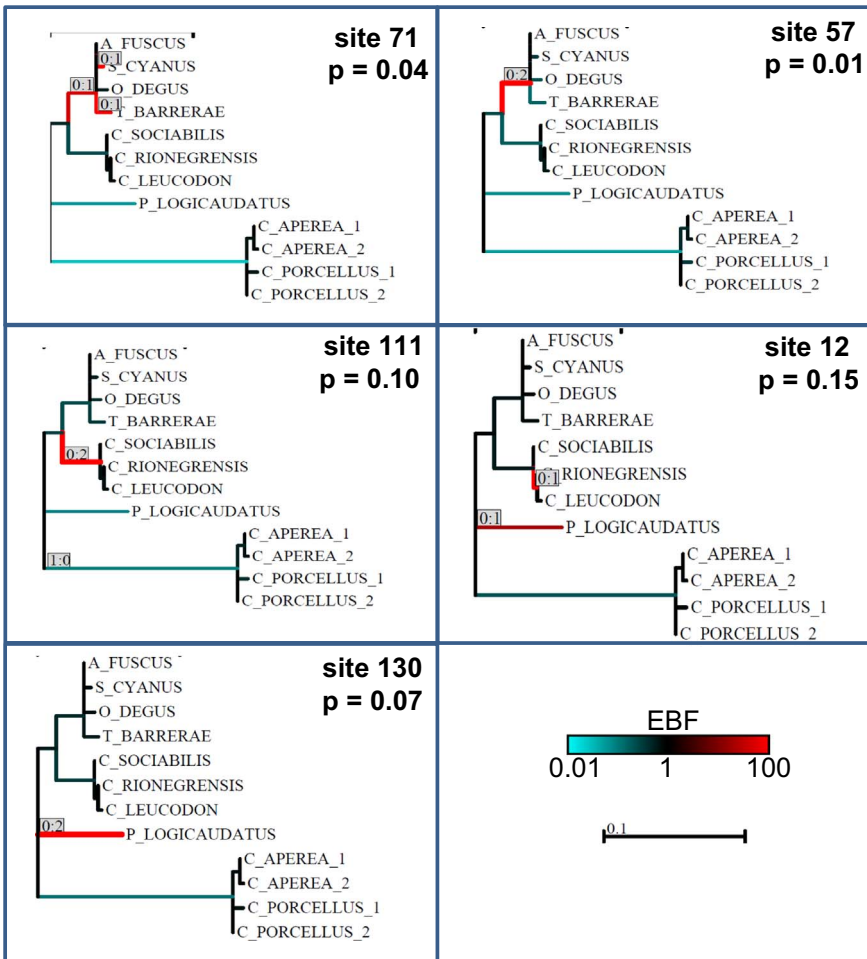

Fig. 2. Summary of results obtained for each site found to be under episodic positive selection with MEME [3]. Codon site and probability are specified. Each branch is labeled with the pair Syn: NonSyn substitution counts. Branches with more total inferred substitutions are also drawn with thicker lines. Each branch is colored according to the magnitude of the Empirical Bayes Factor (EBF) for having $\omega > 1$ along that branch. Brighter reds higher EBF values (also see the scale bar).

Table 1

Significant physicochemical amino acid changes among residues in alpha-globin gene identified by TreeSAAP.

		Properties (*: p<0.01, **:p<0.001)				
		Long-range non-bonded energy	Alpha-helical tendencies	Turn tendencies	Coil tendencies	Power to be at the N-terminal
Codon	Branch					
4	node#16 --> C_sociabilis		6 *		8 *	
4	node#11 --> node#18	1 **	8 **		8 *	7 *
4	node#18 --> C_aperea	1 **	6 *			
5	node#16 --> C_sociabilis	1 **	8 **		8 *	7 *
12	node#16 --> node#17	1 **	6 *			
12	node#11 --> P_longicaudatus	1 **	6 *			
13	node#11 --> node#18	1 **	6 *			
21	node#15 --> S_cyanus	1 **	8 **		8 *	
21	node#11 --> node#18	1 **	8 **		8 *	
22	node#18 --> C_aperea	1 **	8 **		8 *	7 *
23	node#11 --> P_longicaudatus	1 **		6 *		
25	node#11 --> node#18				8 *	
28	node#11 --> node#18	1 **	8 **		8 *	
46	node#11 --> node#18	1 **				
48	node#13 --> T_barrerae	1 **				
50	node#12 --> node#16	1 **	8 **		8 *	7 *
50	node#11 --> node#18	1 **	8 **		8 *	7 *
53	node#11 --> node#18	1 **	8 **		8 *	
54	node#11 --> node#18	1 **				
55	node#12 --> node#16	1 **				
57	node#12 --> node#13	1 **	6 *	6 *		
63	node#11 --> node#12	1 **	6 *	6 *		
64	node#11 --> P_longicaudatus	1 **		6 *		
71	node#15 --> S_cyanus	1 **				
71	node#13 --> T_barrerae	1 **		6 *		
82	node#11 --> node#18	1 **	6 *			
110	node#11 --> node#12	1 **	6 *	6 *		
116	node#12 --> node#13	1 **		6 *		
116	node#11 --> node#18	1 **		6 *		
120	node#11 --> node#18	1 **	6 *	6 *		
121	node#11 --> P_longicaudatus		6 *			
125	node#13 --> T_barrerae	1 **				
130	node#11 --> P_longicaudatus	1 **	6 *	6 *		
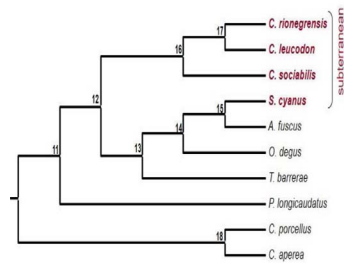
Category 1 is the most conservative; categories from 6 to 8 are radical changes, being the 8 the most radical.
Codons highlighted in gray were suggested to be under positive selection by other methods.

2. Experimental design, materials and methods

We selected 10 caviomorph rodent species, including two independent subterranean lineages (Ctenomys and Spalacopus cyanus), three fossorial but non-subterranean relatives (Aconaemys fuscus, Octodon degus and Tympanoctomys barrerae), a non-fossorial spiny rat (Proechimys longicaudatus), and two species of Cavia (C. porcellus and C. aperea) used as outgroup. We chose a sample of 3 species from Ctenomys in an attempt to capture the diversity of lifestyles, molecular differentiation and habitat features [4].

Total DNA extractions were made from liver preserved in 95% ethyl-alcohol [5]. PCR amplification were carried out with two pair of primers specially designed for this study [1]. PCR products were purified and automatic sequencing from both ends was done by Macrogen. Inc. (http://www.macrogen.com), under BigDyeTM terminator cycling conditions in an ABI 3730xl Sequencer. The intron-exon structure and inspections of the lack of stop codons and in/dels that generate frameshift were checked. Aligned CDS are shown in Fig. 1.

Codon sites with evidence of have evolved under positive selection by different methods has been shown and summarized in [1]. Here we show additional results obtained with MEME [3] implemented in Datamonkey web server (http://www.datamonkey.org/, [6]) and TreeSAAP [2] software. MEME

evaluates variation in dN/dS along branches, and then assess which sites contribute to variation in dN/dS, and the level for significance was 0.05 for the site and, conditional upon that, an Empirical Bayes Factor (EBF) greater than 20 for the branch or branches. TreeSAAP compares physicochemical properties of the amino acid changes observed. Nonsynonymous changes were considered to be the result of positive, destabilizing selection (from now on, "radical changes") only if they met two stringent criteria: a) they were assigned to the most extreme categories of structural or functional changes (categories 6, 7 and 8 of [2]); and b) they were significant at the $p < 0.001$. We ran ModelGenerator [7] to select a substitution model for each gene and reconstruct ancestral sequences in Baseml [8], and the model selected was REV [9]. The phylogeny considered for this study is shown in Fig. 2 and Table 1 and was obtained by pruning the comprehensive tree of octodontoid genera in [10,11] and [4].

Acknowledgements

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2017.09.017.

References

[1] I.H. Tomasco, N. Boullosa, F.G. Hoffmann, E.P. Lessa, Molecular Adaptive Convergence in the α-Globin Gene in Subterranean Octodontid Rodents, 628, 2017, pp. 275–280. http://dx.doi.org/10.1016/j.gene.2017.07.057.
[2] S. Woolley, J. Johnson, M.J. Smith, K.A. Crandall, D.A. McClellan, TreeSAAP: selection on amino acid properties using phylogenetic trees, Bioinformatics 19 (2003) 671–672. http://dx.doi.org/10.1093/bioinformatics/btg043.
[3] B. Murrell, J.O. Wertheim, S. Moola, T. Weighill, K. Scheffler, S.L. Kosakovsky Pond, Detecting individual sites subject to episodic diversifying selection, PLoS Genet. (2012), http://dx.doi.org/10.1371/journal.pgen.1002764.
[4] I.H. Tomasco, E.P. Lessa, The evolution of mitochondrial genomes in subterranean caviomorph rodents: adaptation against a background of purifying selection, Mol. Phylogenet. Evol. 61 (2011) 64–70. http://dx.doi.org/10.1016/j.ympev.2011.06.014.
[5] S.A. Miller, D.D. Dikes, H.H. Polesky, A simple salting procedure for extracting DNA from human nucleated cells, Nucleic Acids Res. 16 (1988) 215.
[6] W. Delport, A.F.Y. Poon, S.D.W. Frost, S.L. Kosakovsky Pond, Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology, Bioinformatics 26 (2010) 2455–2457. http://dx.doi.org/10.1093/bioinformatics/btq429.
[7] T.M. Keane, C.J. Creevey, M.M. Pentony, T.J. Naughton, J.O. McInerney, Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified, BMC Evol. Biol. 6 (2006) 29. http://dx.doi.org/10.1186/1471-2148-6-29.
[8] Z.H. Yang, PAML: a program package for phylogenetic analysis by maximum likelihood, Comput. Appl. Biosci. 13 (1997) 555–556.
[9] C. Saccone, C. Lanave, G. Pesole, G. Preparata, Influence of base composition on quantitative estimates of gene evolution, Methods Enzymol. 19 (1990) 1572–1574.
[10] N.S. Upham, B.D. Patterson, Diversification and biogeography of the Neotropical caviomorph lineage Octodontoidea (Rodentia: hystricognathi), Mol. Phylogenet. Evol. 63 (2012) 417–429. http://dx.doi.org/10.1016/j.ympev.2012.01.020.
[11] A. Parada, G. D'Elía, C.J. Bidau, E.P. Lessa, Species groups and the evolutionary diversification of tuco-tucos, genus Ctenomys (Rodentia: ctenomyidae), J. Mammal. 92 (2011) 671–682. http://dx.doi.org/10.1644/10-MAMM-A-121.1.

Corrigendum

Corrigendum to "Nucleotide sequence and results of test of adaptive evolution in the α-globin gene of octodontoid rodents" [Data Brief 15 (2017) 517–521]

I.H. Tomasco [a,*], N. Boullosa [a], F.G. Hoffmann [b,c], E.P. Lessa [a]

[a] *Departamento de Ecología y Evolución, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo 11400, Uruguay*
[b] *Department of Biochemistry and Molecular Biology, Mississippi State University, MS, USA*
[c] *Institute for Genomics, Biocomputing and Biotechnology, Mississippi State University, MS, USA*

The authors regret that the article's author list is incomplete and author Sophie Tandonnet is missing. The author list should have appeared as follows:

I.H. Tomasco,[a,*] N. Boullosa,[a] S. Tandonnet,[b] F.G. Hoffmann,[c,d] and E.P. Lessa[a].

[a] Departamento de Ecología y Evolución, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo 11400, Uruguay.

[b] School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK.

[c] Department of Biochemistry and Molecular Biology, Mississippi State University, MS, USA.

[d] Institute for Genomics, Biocomputing and Biotechnology, Mississippi State University, MS, USA.

The authors would like to apologise for any inconvenience caused.