

Animal Mitochondria, Positive Selection and Cyto-Nuclear Coevolution: Insights from Pulmonates

Aristeidis Parmakelis^{1*}, Panayiota Kotsakiozi², David Rand³

1 Department of Ecology and Taxonomy, Faculty of Biology, National and Kapodistrian University of Athens, Athens, Greece, **2** Department of Animal and Human Physiology, Faculty of Biology, National and Kapodistrian University of Athens, Athens, Greece, **3** Department of Ecology and Evolutionary Biology, Brown University, Providence, Rhode Island, United States of America

Abstract

Pulmonate snails have remarkably high levels of mtDNA polymorphism within species and divergence between species, making them an interesting group for the study of mutation and selection on mitochondrial genomes. The availability of sequence data from most major lineages – collected largely for studies of phylogeography – provides an opportunity to perform several tests of selection that may provide general insights into the evolutionary forces that have produced this unusual pattern. Several protein coding mtDNA datasets of pulmonates were analyzed towards this direction. Two different methods for the detection of positive selection were used, one based on phylogeny, and the other on the McDonald-Kreitman test. The cyto-nuclear coevolution hypothesis, often implicated to account for the high levels of mtDNA divergence of some organisms, was also addressed by assessing the divergence pattern exhibited by a nuclear gene. The McDonald-Kreitman test indicated multiple signs of positive selection in the mtDNA genes, but was significantly biased when sequence divergence was high. The phylogenetic method identified five mtDNA datasets as affected by positive selection. In the nuclear gene, the McDonald-Kreitman test provided no significant results, whereas the phylogenetic method identified positive selection as likely present. Overall, our findings indicate that: 1) slim support for the cyto-nuclear coevolution hypothesis is present, 2) the elevated rates of mtDNA polymorphisms and divergence in pulmonates do not appear to be due to pervasive positive selection, 3) more stringent tests show that spurious positive selection is uncovered when distant taxa are compared and 4) there are significant examples of positive selection acting in some cases, so it appears that mtDNA evolution in pulmonates can escape from strict deleterious evolution suggested by the Muller's ratchet effect.

Citation: Parmakelis A, Kotsakiozi P, Rand D (2013) Animal Mitochondria, Positive Selection and Cyto-Nuclear Coevolution: Insights from Pulmonates. PLoS ONE 8(4): e61970. doi:10.1371/journal.pone.0061970

Editor: Kostas Bourtzis, International Atomic Energy Agency, Austria

Received: January 30, 2013; **Accepted:** March 15, 2013; **Published:** April 19, 2013

Copyright: © 2013 Parmakelis et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This project was partially funded by the "John S. Latsis" Public Benefit Foundation, and the sole responsibility for the content lies with its authors. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding was received for this study.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: aparmakel@biol.uoa.gr

Introduction

MtDNA loci are the most commonly used markers in molecular studies of animal populations [1]. Studies involving the genetic structure of populations, phylogenetic relationships of animal taxa and the spatial and temporal evolution of populations, often involve mtDNA haplotyping. The reasons for the adoption of mtDNA as the favorable marker for molecular studies are well-known (reviewed in [1]). Among other biological attributes, mtDNA is considered to be evolving in a nearly neutral fashion. Being involved in basic metabolic functions (respiration), mitochondrial-encoded genes have been considered as less likely than other genes to be adaptively evolving. However, the fundamental property of the neutral evolution of the mtDNA genome has been questioned [1].

Numerous cases exist [2,3,4,5,6,7,8,9,10] in favor of the view that the variation of the mitochondrial genome is not selectively neutral. Additionally, positive selection has been implicated in shaping the mitochondrial diversity recorded in several animal species [5,6,7,10,11,12,13,14,15] including humans [11,16], whereas arguments in favor of purifying selection are also

expressed [17]. Fitness consequences [18] are pertinent to the mitochondrial genetic variation that exists within and between populations, and mechanisms through which sequence variation in mtDNA is achieved [19] and maintained [20,21] have been proposed. Therefore, as it has been concluded [22], "the mtDNA genome is a profound player in the evolutionary dynamics of populations, and the existence of non-neutral genetic variation in mtDNA must be a common phenomenon".

Aiming at the investigation of the mode of evolution of the mtDNA genome in animal species, we focused on land-snails, a group of organisms that exhibit extreme levels of mtDNA divergence compared to other animal groups [23,24,25,26]. Several researchers have attempted to assess the relative importance of each of the four not mutually exclusive hypotheses that were proposed [23] in order to explain the extreme intra-specific mtDNA divergence observed in pulmonates [23]. These hypotheses are: (i) mtDNA evolution in pulmonates is exceptionally fast, (ii) haplotypes differentiate in isolated 'refuges' and subsequently come together, (iii) natural selection is acting to preserve variation, and (iv) the population structure of pulmonates favors the persistence of ancient haplotypes. The accelerated mtDNA

evolution hypothesis was supported by the findings in some taxa [24,26,27,28], but not in others [29,30,31], whereas a hypothesis was formulated, on developmental grounds, to account for the extraordinary mtDNA diversity in mollusks [32]. At the same time, several studies concluded that a history of allopatric divergence followed by secondary contact (hypothesis ii) could explain sympatry of diverse haplotypes [24,30,31,33,34,35,36]. The effect of natural selection in shaping the population genetic structure of pulmonates (hypothesis iii) has not been directly investigated, but there were cases where natural selection was considered to have played a major role [33,37,38]. Furthermore, the potential importance of localized adaptations was demonstrated in *Cepaea* mtDNA [6]. Finally, in some studies it was shown that demographic factors (hypothesis iv), such as highly structured populations and limited capacity for active dispersal, are important both in conservation of mtDNA polymorphisms at a small-scale [39,40] and in the long-term and large-scale structuring of populations [28,30,37,41]. Moreover, it has been proposed [23,40] that even in the face of adequate gene flow within demographically structured pulmonate populations, natural selection could favor specific associations between mtDNA and nuclear genes (associations developed whilst populations are separated) and that would radically slow the homogenization process of the mtDNA lineages, thus allowing several divergent mtDNA haplotypes to be maintained within a single pulmonate population.

The hypothesis of natural selection acting on the combined cyto-nuclear genotype, is continuously gaining support [42,43] as one of the mechanisms through which mtDNA diversity is maintained both within and between animal populations. In this hypothesis, the nuclear genes that interact with the mtDNA genes involved in oxidative phosphorylation (OXPHOS), are the ones more likely to be evolving under the effect of positive selection [42]. These genes are posited to be under a selective regime to maintain functionality of mitochondrial enzyme complexes, driven by the continued loss of function in mtDNA-encoded proteins that experience higher deleterious substitution rates due to reduced effective population size and low recombination rates. Given a high mutation rate in mtDNA, nuclear genes that encode interacting proteins may be more likely to provide positive, compensatory mutations that restore function to OXPHOS complexes that experience deleterious mtDNA mutations [42].

The high levels of mtDNA polymorphism and divergence in pulmonates – and the availability of sequence data from most major lineages – provide a powerful dataset for testing several hypotheses about selection in mtDNA. In an effort to evaluate the role of selection in shaping mtDNA diversity of pulmonates, all the protein coding mtDNA sequences of several pulmonates available in GenBank were analyzed along with sequence data newly generated. The effect of natural selection was investigated through i) extensive phylogenetic analyses, and ii) the implementation of the McDonald-Kreitman test [44]. Additionally, in order to investigate the cyto-nuclear coevolution hypothesis that is proposed to account for the high levels of intra- and inter-population mtDNA divergence, we assessed the divergence pattern exhibited by a nuclear gene (part of the OXPHOS), in two land-snail genera. The nuclear gene investigated is cytochrome *c* (*Cyc*), a gene that has been identified previously as exhibiting signs of molecular adaptation [18,45,46,47].

Phylogenetic methods are widely used to estimate the evolutionary rates of genes and genomes and to detect footprints of natural selection [48]. To examine whether selection is acting on any of the mtDNA genes investigated, a phylogenetic method to infer the pattern of selection was applied. This method detects elevated dN/dS ratios (ω) under specific models of codon

substitution, using maximum likelihood approaches. When the sequences of the reconstructed phylogenetic tree evolve under neutrality, the ratio (also known as ω) of non-synonymous (dN) to synonymous (dS) substitutions is expected to be 1. In the case of positive selection, amino acid changes are favored and $\omega > 1$, whereas under purifying selection (i.e. fewer amino acid substitutions than expected by chance) amino acid changes are prevented and $\omega < 1$. A different test widely used to infer positive selection is the McDonald-Kreitman test [44] that compares the dN/dS ratio between species to that within species and is based on the neutral assumption that this ratio is the same for polymorphism and divergence data. Under positive selection the dN/dS ratio should be larger between species due to excess fixation of amino acid substitutions. Under purifying selection the dN/dS ratio should be larger within species due to an excess of mildly deleterious amino acid polymorphisms that do not reach high frequencies in the population.

Materials and Methods

mtDNA Sequence Data Retrieved from GenBank and Generated for this Study

Data deposited in GenBank up to October 2009 were used. An exception to this is a part of the dataset of *Codringtonia* (see Table 1) that we recently deposited in GenBank. Only coding mtDNA gene sequences of pulmonates were used. The criterion for a specific genus or species to be included in the analysis was that at least seven sequences for at least one coding mtDNA gene should be available. These sequences could originate from different species of the genus (genus analysis) or from different populations of a single species (single species analysis). Sequence datasets were compiled according to genus and mtDNA locus, with the help of CodonCode Aligner v.2.0.6 (CodonCode Corporation, Dedham, MA, USA). The genera analyzed per mtDNA gene are presented in Table 1. In some of the published datasets, the authors concluded that based on the molecular data, species revisions were necessary. In these cases extreme caution was taken to allocate each sequence to the species or genus claimed by the authors (but see further below).

Besides the mtDNA data retrieved from GenBank, we also analyzed unpublished mtDNA sequences (two genes) generated for the present study, originating from one additional pulmonate genus. This genus is *Agathylla* (Westerlund, 1898) (Table 1 and Table S1 in File S1). Furthermore, we analyzed several unpublished *COXI* sequences of *Codringtonia* Kobelt, 1898 (Table 1 and Table S1 in File S1) a genus already represented in GenBank records. Moreover, we generated and analyzed (Table 1 and Table S1 in File S1) mtDNA sequences for *Albinaria* Vest, 1867 that is represented in GenBank but with a different set of species. Details regarding DNA extraction methods, PCR amplification and sequencing of the mtDNA fragments in *Agathylla*, *Albinaria* and *Codringtonia* are provided in the supporting information file (File S1). Accession numbers of generated sequences are provided in Table 1 and Table S1 in File S1.

In the *COXI* sequences, in total 37 genera of pulmonates were represented, whereas in the remaining mtDNA genes 1–4 genera were represented. An overview of the genera analyzed is provided and the reference from which the data originate is cited in Table 1.

Basic analyses of polymorphism and divergence were performed using the software DNAsp v.5.10 [49] and MEGA v.4.0 [50]. The estimated parameters were the number of non-synonymous substitutions per non-synonymous site (dN or K_a) and the number of synonymous substitutions per synonymous site (dS or K_s). Two additional parameters estimated were the “net between group

Table 1. The pulmonate genera and genes analyzed under the maximum likelihood approach inference of adaptive evolution using PAML.

Gene	Genus	Number of species	Length of alignment/ Number of congeneric sequences in alignment	Reference/GenBank Accession Numbers	McDonald-Kreitman test (Y/N)
COX1					
	<i>Achatinella</i>	10	675/77	[87,88]	Y
	<i>Agathylla</i>	11	729/29	see supporting information file	Y
	<i>Albinaria1@</i>	9	651/11	[89,90,91]	N
	<i>Albinaria2@</i>	9	800/40	see supporting information file	Y
	<i>Amborhytida</i>	6	948/36	[92]	Y
	<i>Arianta</i>	4	594/64	[91,93]	Y
	<i>Arion*</i>	25/1	549/38	[94,95,96]	Y
	<i>Ashmunella</i>	5	597/42	Unpublished: AY823787-AY823831	Y
	<i>Bulimulus</i>	23	450/87	[97]	Y
	<i>Candidula</i>	7	522/11	[91,98]	Y
	<i>Carinigera</i>	10	627/31	[90,99]	Y
	<i>Codringtonia*</i>	8/1	702/83	[84]; see supporting information file	Y
	<i>Elona**</i>	1	681/36	[100]	Y
	<i>Euchemotrema**</i>	1	648/12	[101]	Y
	<i>Euhadra</i>	15	552/77	[34,102]	Y
	<i>Everettia</i>	14	558/46	[103,104]	Y
	<i>Iberus</i>	9	627/85	[105,106,107]	Y
	<i>Isabellaria</i>	14	651/24	[89,90]	N
	<i>Kovacsia**</i>	1	654/21	[108]	Y
	<i>Leptaxis</i>	5	375/9	[109]	N
	<i>Lozekia**</i>	1	654/17	[108]	Y
	<i>Marmorana</i>	7	597/33	[110]	Y
	<i>Natalina</i>	12	978/62	[111]	Y
	<i>Oreohelix</i>	7	687/35	[112]	Y
	<i>Partula</i>	11	663/146	[113,114,115]	Y
	<i>Partulina</i>	12	675/29	[88]	Y
	<i>Paryphanta**</i>	1	948/15	[92]	Y
	<i>Placostylus</i>	6	672/20	[115,116]	Y
	<i>Prophysaon</i>	8	618/43	[117]	N
	<i>Pupilla</i>	2	654/10	[118]	Y
	<i>Satsuma*</i>	25/1	666/183	[89,90,119]	Y
	<i>Sericata</i>	12	627/24	[89,90]	Y
	<i>Solatopupa</i>	7	342/14	[120]	Y
	<i>Succinea</i>	14	612/61	[115,121]	Y
	<i>Trochulus</i>	10	558/97	[122]	Y
	<i>Vertigo</i>	21	654/65	[123]	Y
	<i>Wainuia</i>	6	414/21	[92,124]	Y
	<i>Xerocrassa</i>	11	633/122	[125]	Y
COX2					
	<i>Agathylla</i>	11	549/29	see supporting information file	Y
	<i>Codringtonia*</i>	8/1	558/82	[84]	Y
	<i>Gnarosophia**</i>	1	480/41	[126]	Y
	<i>Sphaerospira</i>	7	465/58	[126]	Y
Cytb					
	<i>Euhadra</i>	19	726/63	[127,128]	Y
	<i>Partula</i>	19	633/29	[53]	N

Table 1. Cont.

Gene	Genus	Number of species	Length of alignment/ Number of congeneric sequences in alignment	Reference/GenBank Accession Numbers	McDonald-Kreitman test (Y/N)
<i>ATPase 8</i>	<i>Pupilla</i>	2	396/8	[118]	Y
	<i>Albinaria1*</i>	3/1	165/32	[129]	Y
	<i>Cepaea**</i>	1	153/32	[6]	Y
<i>ND1</i>	<i>Arion</i>	20	435/40	[130]	Y
	<i>Euhadra</i>	19	261/63	[127]	Y
<i>ND4L</i>	<i>Euhadra</i>	19	255/63	[127]	Y
<i>Cyc</i>	<i>Albinaria2</i>	9	249/19	see supporting information file	Y
	<i>Codringtonia</i>	6	249/19	see supporting information file	Y

*PAML analyses performed both on multiple species and on a single species (multiple populations per species).

**PAML analyses performed separately on single species of the genus (multiple populations per species).

©The two datasets of *Albinaria* in COX1 include different species.

Length of sequences is given in base pairs (bp). The datasets that were subjected to the McDonald-Kreitman test are indicated with Y. Accession numbers are provided for sequence data generated for this study, and for sequence data available in GenBank but not published.

doi:10.1371/journal.pone.0061970.t001

average” level of sequence divergence between species and the mean intra-species divergence. These parameters were estimated separately for each gene and genus, using MEGA v.4.0 and the Kimura two-parameter distance [51]. For both the mtDNA and the nuclear gene an additional measure of sequence divergence is presented in Table 2. This divergence metric is the tree length (nucleotide substitutions per codon along the tree), and was estimated by Codeml (model M0) of the PAML package [52].

The data used in this work were generated by researchers mostly concerned about the phylogeny of the genus under study. Consequently, it is very possible that even though the majority of the species of a single genus have been included in each study, only a few populations from each species have been used. These conspecific populations could originate from a variety of distant and/or neighboring localities, hence within species level of divergence in some cases might be misleading. Such issues were evaluated on a case-by-case basis and where necessary such cases were excluded from the inference of general patterns.

Land-snail Genera in which the *Cyc* Gene was Amplified and Analyzed

The pulmonate genera in which the nuclear gene (*Cyc*) fragment was amplified are *Codringtonia* and *Albinaria* (Table 1). Details on the geographic origin and species status of the specimens used in the study, as well as accession numbers of generated sequences are given in Tables 1 and S1. Details on specimens’ collection, nucleic acids extraction, cDNA synthesis, RT-PCR, and cloning of the *Cyc* amplicons are provided in the supporting information file.

Investigating Selection Using Maximum Likelihood Analysis (PAML) and the McDonald-Kreitman Test

Extensive description of the alignment of the sequences and the phylogenetic analyses performed, are given in the supporting information file. The generated Bayesian tree of each mtDNA gene and genus (or single species) served as the basis for the implementation of the maximum likelihood methods of the PAML

package [52] aiming at detecting adaptive molecular evolution under specific models of codon substitution. In the vast majority of the cases, outgroups were removed (pruned) from the Bayesian tree prior to the maximum likelihood analysis. In order to evaluate the effect of the pruning scheme applied in the PAML analyses, multiple sets were subjected to a Bayesian analysis without involving outgroup sequences and then were re-analyzed with PAML. The results of both PAML analyses regarding the presence or not of positive selection were identical to each other. Furthermore, the dataset of *Cytb* of *Partula* produced a highly unresolved phylogenetic tree. However, in order to obtain a more representative picture for the evolution of *Cytb*, the dataset of *Partula* was subjected to the PAML analysis using as starting tree the already published one [53]. Finally, in the inference of positive selection for the nuclear *Cyc* gene of *Albinaria* and *Codringtonia*, the *COX1* mtDNA tree of the respective specimens was used. This was done to eliminate the effect of recombination [54] in the inference of positive selection in the nuclear sequence data.

In each one of the genes subjected to the PAML analysis, the likelihood values of the respective phylogenetic tree as being the result of lineages evolving under the assumptions of the site models M0, M3, M1a, M2a, M7 and M8, were estimated. These models (except M0) allow ω (dN/dS) values to vary among different codons. In practice, maximum likelihood estimates that assume a single value of ω (model M0) often produce conservative results because positive selection in a small subset of sites can be confounded by the effects of purifying selection at other sites. A way to circumvent this is to assume heterogeneity among sites and thus evaluate models that allow for variation among sites such as the M1a, M2a, M3, M7 and M8. For a more detailed description of the models used and the approach followed see the supporting information file. In PAML the Naïve empirical Bayes (NEB) [55,56] and the Bayes empirical Bayes (BEB) methods [57] are used to identify sites under positive selection. BEB is implemented under models M2a and M8 only. As suggested in the PAML manual only the results of the BEB method should be considered to be robust in the identification of sites under selection.

Table 2. Likelihood Ratio (LRTs) and McDonald-Kreitman tests' results on the sequence datasets of the mtDNA genes *COX1*, *COX2*, *Cytb*, *ND1*, *ND4L*, *ATP8* and the nuclear gene *Cyc* of pulmonates.

Gene	Pulmonate genus/species	TL	LRT p-value significance in model comparison			Number of amino acids with dN/dS >1	Mc-Donald Kreitman test p-value ^a (npc)	FDR correction for multiple hypothesis testing	Total number of estimable N.I. values
			M0 vs. M3	M1a vs. M2a	M7 vs. M8				
<i>COX1</i>	<i>Achatinella</i>	2.23	**	n.s.	n.s.	—	n.s.		36
	<i>Agathylla</i>	18.4	**	n.s.	n.s.	—	* (1) ** (2)	(3)	9
	<i>Albinaria1</i>	7.45	n.s.	n.s.	n.s.	—	not performed		—
	<i>Albinaria2</i>	10.12	n.s.	n.s.	n.s.	—	n.s.		5
	<i>Amborhytida</i>	17.92	n.s.	n.s.	n.s.	—	n.s.		1
	<i>Arianta</i>	4.84	n.s.	n.s.	n.s.	—	n.s.		51
	<i>Arion</i>	34.29	n.s.	n.s.	n.s.	—	* (4) ** (9)	(9)	89
	<i>Ashmunella</i>	3.11	n.s.	n.s.	n.s.	—	n.s.		6
	<i>Bulimulus</i>	7.33	n.s.	n.s.	n.s.	—	n.s.		119
	<i>Candindula</i>	8.85	n.s.	n.s.	n.s.	—	n.s.		3
	<i>Carinigera</i>	7.48	n.s.	n.s.	n.s.	—	n.s.		8
	<i>Codringtonia</i>	6.5	**	n.s.	n.s.	—	n.s.		12
	<i>Elona quimperiana</i>	0.014	n.s.	n.s.	n.s.	—	n.s.		0
	<i>Euchemotrema hubrichti</i>	0.05	n.s.	n.s.	n.s.	—	n.s.		0
	<i>Euhadra</i>	19.6	n.s.	n.s.	n.s.	—	n.s.		24
	<i>Everettia</i>	12.97	n.s.	n.s.	n.s.	—	n.s.		10
	<i>Iberus</i>	16.03	n.s.	n.s.	n.s.	—	n.s.		55
	<i>Isabellaria</i>	18.58	n.s.	n.s.	n.s.	—	not performed		—
	<i>Kovacsia kovacsii</i>	0.59	n.s.	n.s.	n.s.	—	n.s.		0
	<i>Leptaxis</i>	4.2	n.s.	n.s.	n.s.	—	not performed		—
	<i>Lozekia deubeli</i>	1.35	**	n.s.	n.s.	—	n.s.		15
	<i>Marmorana</i>	5.5	n.s.	n.s.	n.s.	—	n.s.		0
	<i>Natalina</i>	26.03	n.s.	n.s.	n.s.	—	* (7) ** (28)	(31)	55
	<i>Oreohelix</i>	2.34	*	n.s.	n.s.	—	n.s.		12
	<i>Partula</i>	10.72	n.s.	n.s.	**	—	* (1)	(0)	28
	<i>Partulina</i>	1.38	**	n.s.	n.s.	—	n.s.		7
	<i>Paryphanta busbyi</i>	0.18	n.s.	n.s.	n.s.	—	n.s.		1
	<i>Placostylus</i>	14.78	**	n.s.	**	—	n.s.		3
	<i>Prophysaon</i>	11.25	n.s.	n.s.	n.s.	—	not performed		—
	<i>Pupilla</i>	5.55	**	n.s.	n.s.	—	n.s.		1
	<i>Satsuma</i>	30.96	n.s.	n.s.	n.s.	—	* (4)	(0)	88
	<i>Sericata</i>	19.18	n.s.	n.s.	**	—	n.s.		3
<i>Solatopupa</i>	3.71	n.s.	n.s.	n.s.	—	n.s.		2	
<i>Succinea</i>	8.69	n.s.	n.s.	n.s.	—	n.s.		3	
<i>Trochulus</i>	9.56	n.s.	n.s.	n.s.	—	n.s.		63	
<i>Vertigo</i>	14.95	n.s.	n.s.	n.s.	—	n.s.		23	
<i>Wainuia</i>	0.81	*	n.s.	n.s.	—	n.s.		5	
<i>Xerocrassa</i>	13.18	**	n.s.	**	—	* (1)	(0)	39	
<i>COX2</i>	<i>Agathylla</i>	15.44	**	n.s.	n.s.	—	* (4) ** (3)	(5)	15
	<i>Codringtonia</i>	5.96	**	n.s.	n.s.	—	** (1)	(0)	20
	<i>Gnarosophia bellendenkerensis</i>	2.08	**	n.s.	n.s.	—	n.s.		25
	<i>Sphaerospira</i>	12.69	n.s.	n.s.	n.s.	—	n.s.		15
<i>Cytb</i>	<i>Euhadra</i>	15.22	n.s.	n.s.	n.s.	—	* (31) ** (28)	(25)	189

Table 2. Cont.

Gene	Pulmonate genus/species	TL	LRT p-value significance in model comparison			Number of amino acids with dN/dS >1	Mc-Donald Kreitman test p-value ^a (npc)	FDR correction for multiple hypothesis testing	Total number of estimable N.I. values
			M0 vs. M3	M1a vs. M2a	M7 vs. M8				
	<i>Partula</i>	8.74	**	n.s.	*	–	not performed	–	
	<i>Pupilla</i>	0.66	n.s.	n.s.	n.s.	–	n.s.	1	
<i>ND1</i>	<i>Arion</i>	33.83	**	n.s.	n.s.	–	* (12) ** (40) (50)	77	
	<i>Euhadra</i>	19.64	**	n.s.	n.s.	–	* (10) ** (2) (0)	206	
<i>ND4L</i>	<i>Euhadra</i>	18.99	**	n.s.	n.s.	–	* (25) ** (8) (0)	157	
<i>ATP8</i>	<i>Albinaria</i>	2.53	**	n.s.	n.s.	7 ^c	n.s.	4	
	<i>Cepaea nemoralis</i>	1.26	**	n.s.	n.s.	8 ^c	n.s.	12	
<i>Cyc</i>	<i>Albinaria2</i>	0.14	**	*	*	1 ^b , 2 ^c	n.s.	2	
	<i>Codringtonia</i>	0.69	**	n.s.	n.s.	1 ^b	n.s.	5	

*p < 0.05,

**p < 0.01, n.s.: non significant,

^aFisher's test p-value, npc: number of pairwise comparisons that the neutrality index (N.I.) was <1, FDR: False Discovery Rate for multiple-hypothesis-testing correction, ^bindicated (probability >0.99) both by the Bayes empirical Bayes and the Naïve empirical Bayes methods in all model comparisons,^cindicated only by the Naïve empirical Bayes method in all model comparisons.

The tree length (TL) is given as a measure of sequence divergence (nucleotide substitutions per codon along the tree).

doi:10.1371/journal.pone.0061970.t002

In the supporting information file, we are addressing the issue of using recombining sequences in the inference of positive selection through phylogeny. The McDonald-Kreitman test as implemented in DNAsp v.5.10 [49], was applied in all the datasets where feasible (Table 1). In some datasets (e.g. *Cytb* of *Partula*), the McDonald-Kreitman test was not applied, since many of the species were represented by a single specimen. The test was implemented both in pairwise comparisons of different species of a genus and when possible in pairwise comparisons of different populations of a single species. In both cases, the nucleotide sequences of the mtDNA gene under study, were grouped according to their phylogeny (Bayesian tree) and not based on species or population of origin. The neutrality index (N.I.) [58] and the test of Fisher (as implemented in DNAsp v.5.10) were used to evaluate the strength and statistical significance of positive or negative selection indicated by the McDonald-Kreitman test. We considered removing rare single nucleotide polymorphisms (SNPs) from the datasets to see if this would alter the selection patterns inferred by the test. However, we chose not to do so because it is well known that removing low frequency non-synonymous sites will change the McDonald-Kreitman test and N.I. values in the direction of positive selection [59,60]. Our question here is whether a signature of positive selection is pervasive over-and-above the natural level of mildly deleterious variation that is well known to exist for animal mtDNA [61,62]. Therefore, the removal of rare singletons that is applied when trying to discover the proportion of sites that are fixed by positive selection [59], is not justified in our study. For the reasons stated above, we only applied the McDonald-Kreitman test and no modifications or improvements of this test were implemented. When several statistical inferences based on the same dataset are performed simultaneously, hypothesis tests that incorrectly reject the null hypothesis are more likely to occur. To cope with this, a stronger level of evidence is required for an individual test to be stated as significant, to compensate for the number of tests being made. We applied the false discovery rate (FDR) analysis [63] to correct for such pseudoreplicated comparisons by computation of a p-value cut-off

[64] in those datasets for which the McDonald-Kreitman test was in favor of positive selection acting on the sequences.

It has been argued that in datasets where multiple substitutions at each synonymous site have occurred, the number of synonymous substitutions will probably be underestimated and this can lead to erroneous inference of positive selection [61,65,66,67,68]. In order to investigate the issue of substitution saturation argument we estimated (in all datasets) the pairwise estimates [69] of dN and dS and plotted them against the respective pairwise genetic distance calculated based on the Jukes-Cantor model [70]. Relationships between the Jukes-Cantor distance and dN and dS were modeled separately using the ordinary least squares method. The cubic, quadratic and the linear function of the Jukes-Cantor distance were tested in order to account for curvilinear relationships. Details on the model fitting procedure are given in the supporting information file.

Results

Generated *Cyc* and mtDNA Sequences

Details on the generated mtDNA sequences from specimens of *Agathylla*, *Albinaria*, and *Codringtonia* are given in Table S1. We successfully amplified a fragment of the *Cyc* gene in 38 land-snail specimens, 19 from each land-snail genus targeted (Table 1). After assigning the individual specimens to the nominal species they belong to, we had *Cyc* sequences for six *Codringtonia* and nine *Albinaria* species, respectively (Table 1). In all specimens a single *Cyc* allele was present. The length of the obtained fragments ranged from 139 to 202 (average 172 bp) and 59 to 250 bp (average 190 bp) for *Albinaria* and *Codringtonia*, respectively.

Levels of Sequence Divergence between and within Species of Pulmonates

The mean levels of sequence divergence between species (mean inter-specific divergence) and within species (mean intra-specific divergence) for each mtDNA gene and land-snail genus are

presented in Figure S1 in File S1. According to the data available, the mean sequence divergence between different species of pulmonates could be in the range of 2% to 34.5% depending on the mtDNA gene used. At the same time the within species sequence divergence is between 0.2 and 13.8%. A more detailed presentation of the divergence levels is provided in the supporting information file.

Thirty-eight sequences of *Cyc* from two land-snail genera were produced for this study. The between-species levels of sequence divergence for *Cyc* range from 0 to 3.1% in *Albinaria* and 0 to 5.8% in *Codringtonia*. The within-species level of sequence divergence is in the range of 0 to 4.3% in *Codringtonia* and 0 to 0.6% in *Albinaria*.

Detecting Selection Using Maximum Likelihood Analysis and the McDonald-Kreitman Test

All the datasets were subjected to several tests designed to detect recombination (see supporting information file) which was not identified in any of the datasets that provided evidence for positive selection (PAML or McDonald-Kreitman). In Table 2 we are providing the results of the maximum likelihood analysis and the McDonald-Kreitman tests on the available datasets. More detailed results per evaluated gene are presented below.

In total, 41 datasets of *COXI* sequences originating from 37 pulmonate genera (Table 1) were analyzed using maximum likelihood analysis and investigated for the presence of positive selection. Out of the 41 datasets, the LRT in 10 of them (25% of the datasets) indicated that varying selection pressure (M0 versus M3) is acting along the protein (Table 2). However, no amino acids were indicated as evolving under positive selection. The *COXI* datasets were also tested with the McDonald-Kreitman test and in six cases the results were in favor of the presence of positive selection (Table 2). The neutrality index (N.I.) in several pair-wise comparisons within the genera of *Agathylla*, *Arion*, *Natalina*, *Partula*, *Satsuma* and *Xerocrassa* was significantly less than 1 and the p-value of the Fisher's test was significant. After correcting for multiple comparisons using the FDR, no support for positive selection was evident in the datasets of *Partula*, *Satsuma* and *Xerocrassa* based on the McDonald-Kreitman test (Table 2).

When the M7 model was evaluated against the M8 model in PAML, four *COXI* datasets of pulmonates out of the 41 (10% of the datasets) were in favor of the M8 model being more appropriate for the specific data (Table 2). Two of these datasets indicated positive selection in the M0 versus M3 model comparison as well. Once more, even though the presence of positive selection is favored by the model comparison, no amino acids were indicated as exhibiting dN/dS values greater than 1.

The M1a versus M2a model test did not produce any significant results supporting positive selection acting on the *COXI* gene in any of the pulmonate datasets. The same is true for all remaining mtDNA genes and therefore we will not be referring to this model comparison for the remaining mtDNA genes.

Five *COX2* gene datasets were analyzed (Table 1). In the maximum likelihood analysis (PAML) three of them were in favor of the M3 model versus the M0 (Table 2). This signifies that there is variation in the selective pressure acting along the protein sequence. On the other hand, no amino acids were identified as being affected by positive selection. The McDonald-Kreitman test performed on *COX2* datasets, produced several significant results supporting the possible role of positive selection in shaping the molecular divergence of the sequences in two land-snail genera (*Agathylla* and *Codringtonia*). After the FDR correction only the dataset of *Agathylla* was still indicating signs of positive selection (Table 2).

Out of the three available *Cytb* datasets of pulmonates, one (*Partula*) exhibited signs of non-uniform selection pressure among the amino acid sites (M0 versus M3). The same result was obtained from the M7 versus M8 test (Table 2). In both model comparisons however, no specific amino acids were recognized as exhibiting dN/dS values above 1. The *Cytb* dataset for *Partula* could not be subjected to the McDonald-Kreitman test since no intra-species data were available. However, the test was implemented in the remaining two *Cytb* datasets and was in favor of positive selection acting on the *Cytb* of *Euhadra*. There were 59 (significantly supported) species pairwise comparisons that returned a N.I. value that was less than 1. After the multiple hypothesis testing correction, a quite high number of pairwise comparisons supporting positive selection (25) was maintained (Table 2).

In total two datasets of pulmonates were available for *ATP8* and both of them were analyzed with PAML and provided evidence of positive selection acting on the ATP8 protein. The evidence was derived only from the M0 versus M3 model comparison (Table 2) and in both datasets several amino acid sites were identified as exhibiting dN/dS values greater than 1. However, these sites were solely identified by the NEB method and not by the BEB method (Table 2). On the other hand, the McDonald-Kreitman test was not in favor of positive selection acting on the protein in any dataset (Table 2).

Two *ND1* datasets were analyzed. Both of them favored the M3 versus the M0 model (Table 2). Once more non-uniform selection pressure acting on the amino acids was identified. The attempt to identify positively selected sites was not successful (Table 2). At the same time, the McDonald-Kreitman test identified 52 cases of pairwise comparisons (total number of estimable N.I.: 77, Table 2) with statistically strong evidence of positive selection in *Arion* and 12 (out of estimable N.I.: 206) in *Euhadra* (Table 2). After the FDR adjustment the number of pairwise comparisons indicating positive selection in *Arion* was 50 and 0 for *Euhadra*.

A single *ND4L* gene dataset (Table 1) was analyzed and the maximum likelihood analyses were not in favor of positive selection acting on this gene. On the other hand, the McDonald-Kreitman test indicated 33 cases (out of estimable N.I.: 157) of pairwise comparisons with strong evidence of the neutrality index being less than 1. Following the FDR correction, no signs of positive selection could be found in this dataset (Table 2).

Signs of positive selection were detected in the *Cyc* dataset of *Albinaria*. All model comparisons were in favor of positive selection, and both the LRT and the BEB analyses supported this finding. The BEB analyses, in all model comparisons, indicated a single amino acid as evolving under the effect of positive selection. In the dataset of *Codringtonia* the LRT results were not in favor of positive selection acting on *Cyc*. However, the BEB analyses strongly supported the presence of a single amino acid as exhibiting adaptation in all model comparisons.

Mutational Saturation

The results of the three different models fitted to describe the relation of dS and dN with genetic distance (Jukes-Cantor) are presented in the supporting information file. Two detailed tables are provided, one for dS (Table S2 in File S1) and one for dN (Table S3 in File S1) along with the figures (Fig. S2a, S2b, S2c, S3a, S3b, S3c in File S1) of dS and dN versus genetic distance for all 45 land-snail datasets.

In the majority of the datasets for both dS and dN either the cubic or the quadratic model provides a better fit. More specifically, in the dS analyses the cubic model shows a better fit in 30 out of the 45 land-snail datasets, whereas the quadratic is more appropriate in 11 datasets and the linear only in four (Table

S2 in File S1, w_i values). Moreover, in the 40 datasets in which a polynomial model exhibits a significant better fit, there are 15 cases (Table S2 in File S1) in which a declining trend of dS values (as a function of the genetic distance) is predicted (i.e. negative value of terminal slope of the best polynomial model).

In the dN analyses, based on the w_i values it is evident that in 31 out of the 45 datasets, the cubic model has a better fit (Table S3 in File S1). The quadratic is indicated as better fitted in 13 datasets, whereas the linear model only in three. Finally, in these 42 datasets that the dN values are better described as a function of the genetic distance using a polynomial model, there are 12 cases where declining values of dN are predicted (Table S3 in File S1).

The results using the AIC were similar to those using the BIC. There were some cases where a different model was selected by the AIC. More specifically, in seven datasets a more complex model (cubic: 6, quadratic: 1) was selected for dS. Regarding dN there were six datasets for which the more complex cubic model was chosen. However, the substitution trends when different models were preferred remained the same.

Aiming at further investigating the effect of mutational saturation on the inference of selection we compared the tree lengths of datasets showing positive selection in the McDonald-Kreitman tests with those that do not. This comparison was performed using a Mann-Whitney U test (results not shown) and it was clearly indicated that the tree length of the analyzed datasets has a significant impact on the proportion of significant MacDonal-Kreitman tests and this relationship is opposite for datasets showing positive or negative selection. It was evident that the tree length is significantly longer for the datasets showing positive selection than for those that are not showing signs of positive selection, while there is no significant difference between datasets showing negative selection and those that do not deviate from neutrality.

Discussion

Land-snails are known to exhibit high levels of mtDNA divergence [23,24,26]. Figure S1 in File S1 presents the mean levels of inter- and intra-specific sequence divergence based on the Kimura-2p model. According to these data, it is evident that both the intra- and inter-specific levels of divergence are extremely high. Divergence between species is as high as 35%, whereas intra-specific differences rise up to 13.8% (Figure S1 in File S1). The levels of divergence in pulmonates are comparable to those found for homologous sequences within taxa described as genera, families or even orders for other organisms, as has already been stated [23]. Therefore, investigating the mode of evolution of the pulmonate mtDNA is challenging. Given the extremely high level of intra-specific sequence divergence reported here (13.8%, supporting information file), it could be argued that certain mtDNA alleles may have diverged for a long time in allopatry before coming together subsequently and might represent populations that are no longer able to interbreed. Such a situation is likely to produce significant results of the McDonald-Kreitman test in which the N.I. values will be greater than 1 but with polymorphic non-synonymous substitutions occurring at intermediate frequencies in the incriminated population, a situation that is quite different from strong purifying selection in which polymorphic non-synonymous mutations occur at low frequency. However, as we mention in the materials and methods section, the sequences were assigned to phylogenetic clades and not nominal species. This was done to avoid inflating the parameters indicating positive selection when cryptic species are included in such analyses [71]. Furthermore, it has been shown that individuals of a

single population diverging as much as 6.3% in their *16S rDNA* sequences can interbreed and exchange spermatophores in the field [25]. Therefore, it cannot be conclusively claimed that individuals from different populations being separated by mtDNA sequences differing as much as 13.8%, cannot interbreed. On the contrary, there are strong indications that they can, but the question remains whether or not they can produce viable offspring.

Other invertebrate groups have also been found to exhibit extreme levels of mtDNA divergence, such as copepods [72] and hymenoptera [13]. In most cases the small effective population size of the mtDNA, is invoked to explain the faster accumulation of mutations in the mitochondrial genome compared to the nucleus. However, researchers are considering several other factors as contributing to this pattern. For instance lineage-specific attributes such as body size [73], generation time [74], effective population size [68,75], and presence of *Wolbachia* [13], have been considered to affect both the mutation rate exhibited as well as the possibility of a mutation becoming fixed in a population. Pulmonates are ubiquitously hermaphrodites [32] and as such the effective population size of the organelles is in theory twice that of normal (all else being equal), since all individuals in a population can potentially contribute to the next generation [32]. Therefore the smaller effective population size argument is weakened to some degree. As already indicated in the introduction, one of the four competing hypotheses that have been proposed to account for the high levels of intra-specific mtDNA divergence of pulmonates is that of natural selection acting on the mtDNA haplotypes and shaping the genetic structure of pulmonates' populations. Through the approach followed in the present work, we evaluated the role of natural selection in shaping the diversity of the protein-coding mtDNA sequences of pulmonates. Moreover, we took a step further and investigated the cyto-nuclear coevolution hypothesis that has also been proposed to account for the high levels of intra- and inter-population mtDNA divergence recorded in animals [42,43].

The inter-specific levels of divergence estimated for *Cyc* in *Albinaria* (0 to 3.1%), are within the range reported for other invertebrate species. However, in *Codringtonia* the values obtained (0 to 5.8%) are higher. In *Tigriopus californicus* [76] the divergence values between different populations (or species according to Willett and Burton [45]) range from 0 to 3.3%, whereas in *Carabus* [77] the inter-specific divergence levels are in the range of 0 to 3.8%. No data are available for intra-specific comparisons.

Indications of Selection in the mtDNA Genes

Based on the phylogenetic and population genetic analyses performed on the mtDNA sequences datasets, it can be claimed that the signals for positive selection obtained from the two different kinds of analyses are partially congruent. Starting from the McDonald-Kreitman test, it provides several indications of positive selection acting on the mtDNA genes evaluated (Table 2). In a preliminary analysis (results not shown) performed on the same datasets but without implementing the phylogenetic grouping scheme in the analysis (see materials and methods section) the results of the McDonald-Kreitman test were providing additional cases of positive selection. However, we have chosen to define intra- and inter-specific groups based on phylogenetic analyses since according to our view it makes more sense, and furthermore it ensures the escape from the pitfall of cryptic species present within the datasets that may well inflate the parameters indicating positive selection, an issue that was addressed in [71] as well. Additionally, when a correction for multiple hypothesis testing was

applied to the results of the McDonald-Kreitman test, the support of positive selection was further reduced (Table 2).

McDonald-Kreitman Test and Substitution Saturation

In numerous studies [61,66,67,68,78] it has been argued that the McDonald-Kreitman is prone to exhibit false positives of selection when the mutation saturation level is high. The levels of dN and dS saturation were investigated in the datasets analyzed in the present work. For this analysis we have to acknowledge that only a limited number of models were tested to see which one describes better the dN and dS values as a function of the genetic distance. However, the models evaluated possess an important feature and this is their ability to show an accelerating or a declining trend of the fitted variable. Especially the cubic model incorporates both trends and can also reach a plateau. In this sense, we consider that we have selected flexible enough models to identify any saturation trend that might be exhibited by the substitution values. One more issue we have to mention is that not all datasets are informative enough for any model to be fitted (e.g. dN values of *Lozekia deubeli COXI*).

Considering all the above, it can be safely concluded that in 14 datasets the dS values are showing a declining trend (Table S2 in File S1) as genetic distance increases. Among these datasets those of *Codringtonia COXI*, *Partula COXI*, *Codringtonia COX2*, *Euhadra CytB*, *Arion ND1*, *Euhadra ND1* and *Euhadra ND4L* are included. These datasets are part of those datasets for which the McDonald-Kreitman test (no FDR correction) provided multiple positive selection indications (Table 2). At the same time, the datasets showing a declining trend of dN as genetic distance increases are 12 (Table S3 in File S1). Out of these, only two (*Codringtonia COXI*, *Albinaria2 COXI*) provide indications of abrupt saturation trend change when distance increases. The remaining datasets show a marginal declining trend of the dN values with elevating genetic distance. Consequently, dS values at high sequence divergence levels appear to be more saturated than dN values. Therefore, when high sequence divergence leads to substitutional saturation, the false positive outcome of the McDonald-Kreitman test should be more or less anticipated. As evolutionary distance increases, the saturation effect is becoming more pronounced, and is affecting dS to a greater degree than dN. At large divergence values dS is more likely to be underestimated and this will cause the dN/dS value for between species to be larger than expected. In a McDonald-Kreitman test this will bias estimates towards positive selection (N.I. <1.0). The opposite effect should hold for negative selection (underestimating dS makes it harder to get an NI value >1.0). It makes sense that this bias is weaker than the positive selection bias. Therefore, caution is needed in the interpretation of the findings of the McDonald-Kreitman test in those datasets where sequence divergence is very high and an increase of dN and dS with genetic distance is not occurring. Moreover, the positive relationship (results not shown) between tree length (Table 2) and positive MacDonal-Kreitman tests (no FDR correction) is also in favor of significant bias existing in this analysis. This bias results from saturation of synonymous substitutions as has already been claimed in other studies [61,65,66,67,68] but has not been explicitly shown.

Maximum Likelihood Approach for Positive Selection

In the maximum likelihood approach for positive-selection-inference the results provide a different picture. The possibility of our findings being obscured by recombination has been eliminated since the recombination tests performed were negative. Therefore, based on the results of the M0 vs. M3 model comparison (Table 2), it is obvious that regardless of the locus under investigation,

adequate indications exist to safely support the view that the selective pressure exerted along an mtDNA amino acid sequence is varying. Considering that the M0 model is actually an unrealistic one, its rejection is anticipated. Therefore, caution should be exerted when only M0 vs. M3 model indicate positive selection [78]. At the same time, there is not even a single dataset where the M1a vs. M2a model provides indications of positive selection. Consequently, according to this quite conservative test [78], the amino acids of the coding mtDNA sequences of pulmonates evolve under the strict control of purifying selection.

However, in the M7 vs. M8 model comparison (Table 2), some datasets do indicate that the sequence divergence of certain mtDNA genes is affected by positive selection. The datasets that such signs exist are *Partula*, *Placostylus*, *Sericata* and *Xerocrassa* in *COXI* and *Partula* in *Cytb*. Moreover, two of the datasets (*Partula*, *Xerocrassa*) pointed by the M7/M8 model comparison as exhibiting dN/dS >1 are also identified by the McDonald-Kreitman test (no FDR correction). One issue to consider regarding these cases where positive selection based on the M7/M8 LRT is identified is that no amino acids are recognized by the BEB method as being affected by positive selection. The M7/M8 model test is a quite stringent test of positive selection [78] and as such indications of positive selection based on this test should be given serious consideration. The fact that the BEB method cannot pinpoint the amino acids being positively selected is a common result in these analyses. It is probably due to the limited information included in the dataset analyzed to infer the sites. It is easier to decide whether such sites exist than to decide where they are. At this point it should be noted that the results of the present study are validating the findings of Goodacre et al. [6] that performed a maximum likelihood analysis on the *Cepaea nemoralis ATP8* gene and identified some sites as having a dN/dS value above 1. In the present study the same amino acids were identified via the NEB analysis of the M3 model.

Finally, in the dataset of *Albinaria ATP8* several amino acids were also found (NEB analysis, M3) to be exhibiting ω values greater than one. Nevertheless, it has been stated [79] and simulations studies have shown [78], that the power and accuracy of both the LRT and the BEB analysis are affected by the number and length of the sequences analyzed and the strength of positive selection [78,79,80]. For the LRT it has been found that longer sequences exhibit an increased probability of detecting adaptive evolution. At the same time very similar sequences carry little information, causing low power of the LRT [78]. On the other hand, multiple substitutions at the same site, as is the case in highly divergent sequences, result in the loss of information and reduced power in identifying positive selection. The simulation studies [78,80] consider a tree length (a metric of sequence divergence) of 0.1 to 1 indicative of low divergence, 1.1 to 10 indicative of medium divergence, whereas values above 10 are considered to be of high divergence. In the present study there is only one case (*Wannia*) where the tree length (Table 2) fits within the low divergence range among the datasets that have shown some signs of positive selection based on the maximum likelihood approaches (Table 2). The remaining datasets are exhibiting medium to high sequence divergence with the majority of them (11/23) leaning towards medium. Regarding the number of sequences involved in our analyses, in all the datasets the sequences used are well above the limits set in the simulation studies [78,80]. Finally, the length of the coding sequences analyzed in the datasets reported in Table 2, is well above the 100 amino acids minimum length required [79,80] for these analyses. Exceptions to this are the *ATP8* gene datasets of *Albinaria* and *Cepaea nemoralis* (Table 1). Therefore, for the datasets of pulmonates that have shown signs of

positive selection based on the M7/M8 LRT, there is no reason to question this finding (Table 2). At this point we have to note that *ATP8* is a gene consistently identified in other animal groups as evolving (even after correction for multiple substitutions, [10,13]) under the effect of positive selection [10,13,15]. In the datasets of *ATP8* of *Albinaria* and *Cepaea* the NEB analysis identified specific sites as exhibiting values of dN/dS >1. *ATP8* could be considered as evolving under the effect of positive selection, however we argue that it is not since: a) it is only the NEB analysis indicating positive selection and b) the length of this particular gene is below the 100 amino acids limit set by simulation studies.

Summarizing the findings from the mtDNA genes, a significantly high number (4/37 or 10.8%) of land-snail genera exhibit signs of positive selection (Table 2, M7/M8 model comparison) in *COX1* gene and one out of two in *Cytb*. In the remaining mtDNA genes no signs of positive selection were detected, however it has to be noted that the number of land-snail genera investigated in all remaining mtDNA genes was significantly smaller than those approached in *COX1*. There was a single land-snail genus (*Partula*) in which both *COX1* and *Cytb* were found to be evolving under the effect of diversifying selection.

Two recent examples of possible adaptive evolution of mtDNA were associated with dramatic metabolic differences in the habits of snakes and bats [10,15]. In these studies the branch-site models test [81] investigating for episodic positive selection, was implemented. In pulmonates, it is possible that the patterns of aestivation in summer and hibernation in winter reflect changes in oxygen utilization that might drive strong diversification. However, these traits are not sufficiently lineage-specific that they can account for the lineage specific patterns of mtDNA variation, so it is difficult to associate the pulmonate mtDNA evolution to specific life history traits.

Indications of Selection in *Cyc*

In *Cyc* the McDonald-Kreitman test did not indicate any signs of positive selection in either of the datasets analyzed (Table 2). On the other hand, the maximum likelihood approach was in favor of positive selection shaping the diversity of this gene in *Albinaria*, and there was a consensus among all the pairwise model comparisons (Table 2). Furthermore, the BEB analysis in all model comparisons of *Albinaria Cyc*, indicated the presence of a single amino acid as exhibiting an ω value greater than 1 with a probability >0.99. In *Codringtonia Cyc* the LRTs did not reveal any signs of positive selection, however the BEB analysis identified a single amino acid (probability >0.99) as having a dN/dS value that exceeds 1. Although the LRT test of M7 versus M8 in *Codringtonia Cyc* (Table 2) was not statistically significant ($p_{\text{value}} = 0.092$), it was in favor of positive selection acting on the gene.

Given the close association between nuclear-encoded and mitochondrial-encoded proteins of the electron transport system (ETS) and the crucial role of the ETS in energy production, the co-adaptation between nuclear-encoded and mitochondrial-encoded proteins has been strongly selected during the evolutionary history of these complexes [45]. According to Willett and Burton [45] two different scenarios of selection/drift could generate intergenomic coevolution: (1) Selection pressure on ETS enzymes could be generated by the adaptation of a species (or population) to its environment leading to changes in both nuclear-encoded and mitochondrial-encoded subunits of ETS complexes. (2) High mutation rates in mtDNA, especially when coupled with fluctuating population sizes, may lead to fixation of deleterious amino acid substitutions and provide selection pressure for subsequent compensatory changes in nuclear genes, a two-step process that could be the driving force of coevolution in ETS

enzymes. One of the nuclear-encoded ETS genes that has repeatedly been found to be consistent with mtDNA/nuclear coevolution is the Cytochrome c (*Cyc*) gene [43,45,82]. The signs of selection, identified by the maximum likelihood approach (Table 2), in the *Cyc* gene of *Albinaria* and *Codringtonia* lend slim support to the cyto-nuclear coevolution hypothesis analyzed above, since it is based upon a single adaptive amino acid substitution. The amplification of the whole *Cyc* gene in the two studied genera, and the inclusion of additional land-snail genera could shed more light on this intriguing finding.

Cyto-nuclear Coevolution

Sequence polymorphism surveys among *Tigriopus californicus* populations show evidence for positive selection at *Cyc* but negative selection at the interacting mitochondrial genes [45]. The same pattern is revealed here in *Cyc* of *Albinaria* and *Codringtonia* and one of its mtDNA-encoded counterparts, *COX1*. This pattern, as was the case in *Tigriopus californicus*, is consistent with the model of asymmetric compensatory coevolution discussed in Rand et al. [42]. At this point we have to note that the selection pattern of *Cyc* in land-snails is similar to that of *Tigriopus*, however it is not recorded between conspecific populations of land-snails. Instead, it is occurring between congeneric species in both land-snail genera that were assessed for *Cyc* polymorphism. However, one cannot rule out the possibility of this pattern being evident between conspecific land-snail populations as well. Such an investigation was not performed in the present study. The smaller spatial or taxonomical level that positive selection can be detected is an issue that needs to be further explored.

It is important to note here that the two land-snail genera, *Albinaria* and *Codringtonia*, are in the two extreme ends of land-snails' diversity. *Albinaria* is a highly speciose genus with more than 200 taxa distributed in Greece [83], whereas *Codringtonia* is an endemic land-snail genus with six species distributed in mainland Greece [84]. At the same time premating isolation mechanisms are in effect non-existent in *Albinaria*, thus hybrids are very frequent [85]. On the other hand, hybridization has not been reported in *Codringtonia*. The fact that positive selection acting on a nuclear-encoded gene *Cyc* that interacts with mtDNA encoded counterparts is primarily detected in *Albinaria*, is in favor of the hypothesis of cyto-nuclear incompatibility acting as a post-mating isolation mechanism in *Albinaria*. This in turn could mean that the interactions between nuclear and mitochondrial genomes represent an important component of hybrid breakdown and species formation in land-snails as has already been suggested for other invertebrates such as *Tigriopus californicus* [86].

Conclusions

According to the phylogenetic and population genetic analyses performed on the mtDNA sequences datasets, it can be said that the signals for positive selection obtained from the two different kinds of analyses are partially congruent. The McDonald-Kreitman test provided several indications of positive selection acting on the mtDNA genes evaluated. However, a thorough analysis of the substitution patterns exhibited by the mitochondrial datasets, and the identification of a strong positive relationship between tree length and positive MacDonal-Kreitman tests (no FDR correction), revealed that this test is biased towards identifying positive selection when sequence divergence is high and mutational saturation is inevitable. At the same time, the phylogenetic inference of positive selection refuted most of the McDonald-Kreitman tests findings and identified four datasets of *COX1* and one of *Cytb* as showing signs of positive selection. In the nuclear gene datasets McDonald-Kreitman tests did not produce any

significant results, whereas the phylogenetic methods indicated that in both land snail genera approached, signs of positive selection are present. Overall, our findings indicate that: 1) slim support for the cyto-nuclear co-evolution is present, 2) the elevated rates of mtDNA polymorphisms and divergence in pulmonates do not appear to be due to pervasive positive selection, 3) more stringent tests show that spurious positive selection is uncovered when distant taxa are compared and 4) there are significant examples of positive selection acting in some cases, so it appears that mtDNA evolution in pulmonates can escape from strict deleterious evolution suggested by the Muller's ratchet effect. The strict prediction of accumulation of deleterious mutations in non-recombining genomes is not upheld in land-snails, given a few cases of positive selection detected. It may be that mtDNA experiences occasional recombination permitting episodes of positive selection, or that occasional beneficial mutations can out-weigh the deleterious consequences of all linked polymorphisms combined. Resolution of this issue requires mechanistic studies that are beyond the scope of this report. Nevertheless, it can be claimed that negative selection does not seem to be an obligatory signature of the mtDNA evolution of pulmonates.

Supporting Information

File S1 Supporting figures and tables. Figure S1. Mean levels of sequence divergence between species (mean inter-species divergence) and within species (mean intra-species divergence) based on the Kimura-2p genetic distance model. The+symbol denotes genera with data limitations that did not permit the calculation of a valid estimate, whereas++denotes genera for which only sequences of a single species were available. Figure S2a. The dS values as a function of the genetic distance using three different models. The best fitted model is in bold line. Blue color corresponds to the linear model, red to the quadratic and green to the cubic model.

References

- Galtier N, Nabholz B, Glemin S, Hurst GDD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol Ecol* 18: 4541–4550.
- Ballard JWO, Rand DM (2005) The population biology of mitochondrial DNA and its phylogenetic implications. *Annu Rev Ecol Evol Syst* 36: 621–642.
- Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Mol Ecol* 13: 729–744.
- Ballard JWO, Kreitman M (1995) Is mitochondrial-DNA a strictly neutral marker? *Trends Ecol Evol* 10: 485–488.
- Blier PU, Dufresne F, Burton RS (2001) Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation. *Trends Genet* 17: 400–406.
- Goodacre SL, Thomaz D, Davies EK (2006) Mitochondrial gene diversity in *Cepaea* - population structure, history and positive selection. *Biol J Linn Soc* 87: 167–184.
- Meiklejohn CD, Montooth KL, Rand DM (2007) Positive and negative selection on the mitochondrial genome. *Trends Genet* 23: 259–263.
- Rand DM (2001) The units of selection on mitochondrial DNA. *Annu Rev Ecol Syst* 32: 415–448.
- Castoe TA, de Koning APJ, Kim H-M, Gu W, Noonan BP, et al. (2009) Evidence for an ancient adaptive episode of convergent molecular evolution. *Proc Natl Acad Sci U S A* 106: 8986–8991.
- Castoe TA, Jiang ZJ, Gu W, Wang ZO, Pollock DD (2008) Adaptive evolution and functional redesign of core metabolic proteins in snakes. *PLoS ONE* 3: e2201.
- Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, et al. (2003) Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci U S A* 100: 171–176.
- Schmidt TR, Wu W, Goodman M, Grossman LI (2001) Evolution of nuclear- and mitochondrial-encoded subunit interaction in cytochrome c oxidase. *Mol Biol Evol* 18: 563–569.
- Oliveira DCSG, Raychoudhury R, Lavrov DV, Werren JH (2008) Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). *Mol Biol Evol* 25: 2167–2180.
- Neiman M, Hehman G, Miller JT, Logsdon JM, Taylor DR (2010) Accelerated mutation accumulation in asexual lineages of a freshwater snail. *Mol Biol Evol* 27: 954–963.
- Shen YY, Liang L, Zhu ZH, Zhou WP, Irwin DM, et al. (2010) Adaptive evolution of energy metabolism genes and the origin of flight in bats. *Proc Natl Acad Sci U S A* 107: 8666–8671.
- Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC (2004) Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* 303: 223–226.
- Kerr KCR (2011) Searching for evidence of selection in avian DNA barcodes. *Mol Ecol Resour* 11: 1045–1055.
- James AC, Ballard JWO (2003) Mitochondrial genotype affects fitness in *Drosophila simulans*. *Genetics* 164: 187–194.
- Neiman M, Taylor DR (2009) The causes of mutation accumulation in mitochondrial genomes. *Proc R Soc Biol Sci Ser B* 276: 1201–1209.
- Dowling DK, Friberg U, Hailer F, Arnqvist G (2007) Intergenic epistasis for fitness: Within-population interactions between cytoplasmic and nuclear genes in *Drosophila melanogaster*. *Genetics* 175: 235–244.
- Rand DM, Clark AG, Kann LM (2001) Sexually antagonistic cytonuclear fitness interactions in *Drosophila melanogaster*. *Genetics* 159: 173–187.
- Dowling DK, Friberg U, Lindell J (2008) Evolutionary implications of non-neutral mitochondrial genetic variation. *Trends Ecol Evol* 23: 546–554.
- Thomaz D, Guiller A, Clarke B (1996) Extreme divergence of mitochondrial DNA within species of pulmonate land snails. *Proceedings of the Royal Society of London Series B-Biological Sciences* 263: 363–368.
- Pinceel J, Jordaens K, Backeljau T (2005) Extreme mtDNA divergences in a terrestrial slug (Gastropoda, Pulmonata, Arionidae): accelerated evolution, allopatric divergence and secondary contact. *J Evol Biol* 18: 1264–1280.
- Parmakelis A, Mylonas M (2004) Dispersal and population structure of two sympatric species of the mediterranean land snail genus *Mastus* (Gastropoda, Pulmonata, Enidae). *Biol J Linn Soc* 83: 131–144.
- Chiba S (1999) Accelerated evolution of land snails *Mandarina* in the oceanic Bonin Islands: evidence from mitochondrial DNA sequences. *Evolution* 53: 460–471.
- Hayashi M, Chiba S (2000) Intraspecific diversity of mitochondrial DNA in the land snail *Euhadra peliomphala* (Bradybaenidae). *Biol J Linn Soc* 70: 391–401.

to the cubic model. Figure S2b. The dS values as a function of the genetic distance using three different models. The best fitted model is in bold line. Blue color corresponds to the linear model, red to the quadratic and green to the cubic model. Figure S2c. The dS values as a function of the genetic distance using three different models. The best fitted model is in bold line. Blue color corresponds to the linear model, red to the quadratic and green to the cubic model. Figure S3a. The dN values as a function of the genetic distance using three different models. The best fitted model is in bold line. Blue color corresponds to the linear model, red to the quadratic and green to the cubic model. Figure S3b. The dN values as a function of the genetic distance using three different models. The best fitted model is in bold line. Blue color corresponds to the linear model, red to the quadratic and green to the cubic model. Figure S3c. The dN values as a function of the genetic distance using three different models. The best fitted model is in bold line. Blue color corresponds to the linear model, red to the quadratic and green to the cubic model.

(PDF)

Acknowledgments

We are grateful to Sinos Giokas for providing the *Albinaria* specimens, and to Francois Rigal that enabled us to perform the model fitting and other statistical analyses. Thanks are due to Zoltan Feher and Zoltan Eross for providing the *Agathylla* specimens. AP would like to thank Spyros Sfenthourakis for linguistic corrections on a previous version of the manuscript.

Author Contributions

Conceived and designed the experiments: AP. Performed the experiments: AP PK. Analyzed the data: AP DR. Contributed reagents/materials/analysis tools: AP PK DR. Wrote the paper: AP PK DR.

28. Thacker RW, Hadfield MG (2000) Mitochondrial phylogeny of extant Hawaiian tree snails (Achatinellinae). *Mol Phylogenet Evol* 16: 263–270.
29. Douris V, Cameron RAD, Rodakis GC, Lecanidou R (1998) Mitochondrial phylogeography of the land snail *Albinaria* in Crete: Long-term geological and short-term vicariance effects. *Evolution* 52: 116–125.
30. Guiller A, Coutellec-Vreto MA, Madec L, Deunff J (2001) Evolutionary history of the land snail *Helix aspersa* in the Western Mediterranean: preliminary results inferred from mitochondrial DNA sequences. *Mol Ecol* 10: 81–87.
31. Pfenninger M, Posada D (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): Fragmentation, corridor migration, and secondary contact. *Evolution* 56: 1776–1788.
32. Davison A (2006) The ovotestis: an underdeveloped organ of evolution. *BioEssays* 28: 642–650.
33. Davison A, Clarke B (2000) History or current selection? A molecular analysis of 'area effects' in the land snail *Cepaea nemoralis*. *Proc R Soc Biol Sci Ser B* 267: 1399–1405.
34. Shimizu Y, Ueshima R (2000) Historical biogeography and interspecific mtDNA introgression in *Euhadra peliomphala* (the Japanese land snail). *Heredity* 85: 84–96.
35. Teshima H, Davison A, Kuwahara Y, Yokoyama J, Chiba S, et al. (2003) The evolution of extreme shell shape variation in the land snail *Ainohelix editha*: a phylogeny and hybrid zone analysis. *Mol Ecol* 12: 1869–1878.
36. Parmakelis A, Pfenninger M, Spanos L, Papagiannakis G, Louis C, et al. (2005) Inference of a radiation in *Mastus* (gastropoda, pulmonata, enidae) on the island of Crete. *Evolution* 59: 991–1005.
37. Davison A (2002) Land snails as a model to understand the role of history and selection in the origins of biodiversity. *Popul Ecol* 44: 129–136.
38. Goodacre SL (2002) Population structure, history and gene flow in a group of closely related land snails: genetic variation in *Partula* from the Society Islands of the Pacific. *Mol Ecol* 11: 55–68.
39. Davison A (2000) The inheritance of divergent mitochondria in the land snail, *Cepaea nemoralis*. *J Molluscan Stud* 66: 143–147.
40. Goodacre SL (2001) Genetic variation in a Pacific Island land snail: population history versus current drift and selection. *Proc R Soc Biol Sci Ser B* 268: 121–126.
41. Sauer J, Hausdorf B (2010) Reconstructing the evolutionary history of the radiation of the land snail genus *Xerocrassa* on Crete based on mitochondrial sequences and AFLP markers. *BMC Evol Biol* 10: 299.
42. Rand DM, Haney RA, Fry AJ (2004) Cytonuclear coevolution: the genomics of cooperation. *Trends Ecol Evol* 19: 645–653.
43. Gerber AS, Loggins R, Kumar S, Dowling TE (2001) Does nonneutral evolution shape observed patterns of DNA variation on animal mitochondrial genomes? *Annu Rev Genet* 35: 539–566.
44. McDonald JH, Kreitman M (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* 351: 652–654.
45. Willett CS, Burton RS (2004) Evolution of interacting proteins in the mitochondrial electron transport system in a marine copepod. *Mol Biol Evol* 21: 443–453.
46. Harrison JS, Burton RS (2006) Tracing hybrid Incompatibilities to single amino acid substitutions. *Mol Biol Evol* 23: 559–564.
47. Burton RS, Ellison CK, Harrison JS (2006) The sorry state of F-2 hybrids: Consequences of rapid mitochondrial DNA evolution in allopatric populations. *Am Nat* 168: S14–S24.
48. Yang ZH (2005) The power of phylogenetic comparison in revealing protein function. *Proc Natl Acad Sci U S A* 102: 3179–3180.
49. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
50. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596–1599.
51. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences. *J Mol Evol* 16: 111–120.
52. Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24: 1586–1591.
53. Goodacre SL, Wade CM (2001) Patterns of genetic variation in Pacific island land snails: the distribution of cytochrome b lineages among Society Island *Partula*. *Biol J Linn Soc* 73: 131–138.
54. Anisimova M, Nielsen R, Yang ZH (2003) Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. *Genetics* 164: 1229–1236.
55. Nielsen R, Yang ZH (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 148: 929–936.
56. Yang ZH, Swanson WJ, Vacquier VD (2000) Maximum-likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites. *Mol Biol Evol* 17: 1446–1455.
57. Yang ZH, Wong WSW, Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol* 22: 1107–1118.
58. Rand DM, Kann LM (1996) Excess amino acid polymorphism in mitochondrial DNA: Contrasts among genes from *Drosophila*, mice, and humans. *Mol Biol Evol* 13: 735–748.
59. Fay JC, Wu CI (2000) Hitchhiking under positive Darwinian selection. *Genetics* 155: 1405–1413.
60. Shapiro JA, Huang W, Zhang CH, Hubisz MJ, Lu J, et al. (2007) Adaptive genetic evolution in the *Drosophila* genomes. *Proc Natl Acad Sci U S A* 104: 2271–2276.
61. Rand DM, Kann LM (1998) Mutation and selection at silent and replacement sites in the evolution of animal mitochondrial DNA. *Genetica* 102–3: 393–407.
62. Nachman MW (1998) Deleterious mutations in animal mitochondrial DNA. *Genetica* 102–103: 61–69.
63. Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29: 1165–1188.
64. Storey JD (2002) A direct approach to false discovery rates. *J R Stat Soc Series B Stat Methodol* 64: 479–498.
65. Obbard DJ, Linton YM, Jiggins FM, Yan G, Little TJ (2007) Population genetics of *Plasmodium* resistance genes in *Anopheles gambiae*: no evidence for strong selection. *Mol Ecol* 16: 3497–3510.
66. Wares JP, Barber PH, Ross-Ibarra J, Sotka EE, Toonen RJ (2006) Mitochondrial DNA and population size. *Science* 314: 1388–1389.
67. Wayne ML, Simonsen KL (1998) Statistical tests of neutrality in the age of weak selection. *Trends Ecol Evol* 13: 236–240.
68. Piganeau G, Eyre-Walker A (2009) Evidence for Variation in the Effective Population Size of Animal Mitochondrial DNA. *PLoS ONE* 4: e4396.
69. Nei M, Gojori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3: 418–426.
70. Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN, editor. *Mammalian Protein Metabolism*. New York: Academic Press. 21–132.
71. Wares JP (2010) Natural distributions of mitochondrial sequence diversity support new null hypotheses. *Evolution* 64: 1136–1142.
72. Rawson PD, Burton RS (2006) Molecular evolution at the cytochrome oxidase subunit 2 gene among divergent populations of the intertidal copepod, *Tigriopus californicus*. *J Mol Evol* 62: 753–764.
73. Thomas JA, Welch JJ, Woolfit M, Bromham L (2006) There is no universal molecular clock for invertebrates, but rate variation does not scale with body size. *Proc Natl Acad Sci U S A* 103: 7366–7371.
74. Thomas JA, Welch JJ, Lanfear R, Bromham L (2010) A generation time effect on the rate of molecular evolution in invertebrates. *Mol Biol Evol* 27: 1173–1180.
75. Bazin E, Glemin S, Galtier N (2006) Population size does not influence mitochondrial genetic diversity in animals. *Science* 312: 570–572.
76. Rawson PD, Brazeau DA, Burton RS (2000) Isolation and characterization of cytochrome c from the marine copepod *Tigriopus californicus*. *Gene* 248: 15–22.
77. Sota T, Vogler AP (2003) Reconstructing species phylogeny of the carabid beetles *Omohopterus* using multiple nuclear DNA sequences: heterogeneous information content and the performance of simultaneous analyses. *Mol Phylogenet Evol* 26: 139–154.
78. Anisimova M, Bielawski JP, Yang ZH (2001) Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol Biol Evol* 18: 1585–1592.
79. Yang ZH, Bielawski JP (2000) Statistical methods for detecting molecular adaptation. *Trends Ecol Evol* 15: 496–503.
80. Anisimova M, Bielawski JP, Yang ZH (2002) Accuracy and power of Bayes prediction of amino acid sites under positive selection. *Mol Biol Evol* 19: 950–958.
81. Zhang JZ, Nielsen R, Yang ZH (2005) Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol* 22: 2472–2479.
82. Rawson PD, Burton RS (2002) Functional coadaptation between cytochrome c and cytochrome c oxidase within allopatric populations of a marine copepod. *Proc Natl Acad Sci U S A* 99: 12955–12958.
83. Giokas S (2000) Congruence and conflict in *Albinaria* (Gastropoda, Clausiliidae). A review of morphological and molecular phylogenetic approaches. *Belg J Zool* 130: 95–103.
84. Kotsakiozi P, Parmakelis A, Giokas S, Papanikolaou I, Valakos ED (2012) Mitochondrial phylogeny and biogeographic history of the Greek endemic land-snail genus *Codringtonia* Kobelt 1898 (Gastropoda, Pulmonata, Helicidae). *Mol Phylogenet Evol* 62: 681–692.
85. Giokas S, Mylonas M, Rolan-Alvarez E (2006) Disassociation between weak sexual isolation and genetic divergence in a hermaphroditic land snail and implications about chirality. *J Evol Biol* 19: 1631–1640.
86. Ellison CK, Burton RS (2008) Interpopulation hybrid breakdown maps to the mitochondrial genome. *Evolution* 62: 631–638.
87. Holland BS, Hadfield MG (2002) Islands within an island: phylogeography and conservation genetics of the endangered Hawaiian tree snail *Achatinella mustelina*. *Mol Ecol* 11: 365–375.
88. Holland BS, Hadfield MG (2004) Origin and diversification of the endemic Hawaiian tree snails (Achatinellidae: Achatinellinae) based on molecular evidence. *Mol Phylogenet Evol* 32: 588–600.
89. de Weerd DRU, Groenenberg DSJ, Schilthuizen M, Gittenberger E (2006) Reproductive character displacement by inversion of coiling in clausiliid snails (Gastropoda, Pulmonata). *Biol J Linn Soc* 88: 155–164.
90. de Weerd DRU, Piel WH, Gittenberger E (2004) Widespread polyphyly among *Alopinia* snail genera: when phylogeny mirrors biogeography more closely than morphology. *Mol Phylogenet Evol* 33: 533–548.

91. Steinke D, Albrecht C, Pfenninger M (2004) Molecular phylogeny and character evolution in the Western Palaearctic Helicidae s.l. (Gastropoda : Stylommatophora). *Mol Phylogenet Evol* 32: 724–734.
92. Spencer HG, Brook FJ, Kennedy M (2006) Phylogeography of Kauri snails and their allies from Northland, New Zealand (Mollusca: Gastropoda: Rhytididae: Paryphantinae). *Mol Phylogenet Evol* 38: 835–842.
93. Gittenberger E, Piel WH, Groenenberg DSJ (2004) The Pleistocene glaciations and the evolutionary history of the polytypic snail species *Arianta arbutorum* (Gastropoda, Pulmonata, Helicidae). *Mol Phylogenet Evol* 30: 64–73.
94. Dvorak L, Backeljau T, Reischuetz PL, Horsak M, Breugelmans K, et al. (2006) *Arion alpinus* Pollonera, 1887 in the Czech Republic (Gastropoda: Arionidae). *Malacol Bohemoslov* 5: 51–55.
95. Pinceel J, Jordaens K, Pfenninger M, Backeljau T (2005) Rangewide phylogeography of a terrestrial slug in Europe: evidence for Alpine refugia and rapid colonization after the Pleistocene glaciations. *Mol Ecol* 14: 1133–1150.
96. Soroka M, Kozłowski J, Wiktor A, Kalusi T (2009) Distribution and genetic diversity of the terrestrial slugs *Arion lusitanicus* Mabille, 1868 and *Arion rufus* (Linnaeus, 1758) in Poland based on mitochondrial DNA. *Folia Biol (Krakow)* 57: 71–81.
97. Parent CE, Crespi BJ (2006) Sequential colonization and diversification of Galapagos endemic land snail genus *Bulimulus* (Gastropoda, Stylommatophora). *Evolution* 60: 2311–2328.
98. Pfenninger M, Posada D, Magnin F (2003) Evidence for survival of Pleistocene climatic changes in Northern refugia by the land snail *Trochoidea geyeri* (Soos 1926) (Helicellinae, Stylommatophora). *BMC Evol Biol* 3: 8.
99. de Weerd DRU, Schneider D, Gittenberger E (2005) The provenance of the Greek land snail species *Isabellaria pharsalica*: molecular evidence of recent passive long-distance dispersal. *J Biogeogr* 32: 1571–1581.
100. Vialatte A, Guiller A, Bellido A, Madec L (2008) Phylogeography and historical demography of the Lusitanian snail *Elona quimperiana* reveal survival in unexpected separate glacial refugia. *BMC Evol Biol* 8: 339.
101. Anderson FE (2007) Population genetics of the carinate pillsnail, *Euchemotrema hubrichti*: genetic structure on a small spatial scale. *Conserv Genet* 8: 965–975.
102. Davison A, Chiba S, Barton NH, Clarke B (2005) Speciation and gene flow between snails of opposite chirality. *PLoS Biol* 3: 1559–1571.
103. Hyman IT, Ho SYW, Jermini LS (2007) Molecular phylogeny of Australian *Helicariionidae*, *Euconulidae* and related groups (Gastropoda : Pulmonata : Stylommatophora) based on mitochondrial DNA. *Mol Phylogenet Evol* 45: 792–812.
104. Liew TS, Schilthuizen M, Vermeulen JJ (2009) Systematic revision of the genus *Everettia* Godwin-Austen, 1891 (Mollusca: Gastropoda: Dyakiidae) in Sabah, northern Borneo. *Zool J Linn Soc* 157: 515–550.
105. Elejalde MA, Madeira MJ, Arrebola JR, Munoz B, Gomez-Moliner BJ (2008) Molecular phylogeny, taxonomy and evolution of the land snail genus *Iberus* (Pulmonata: Helicidae). *J Zool Syst Evol Res* 46: 193–202.
106. Elejalde MA, Madeira MJ, Munoz B, Arrebola JR, Gomez-Moliner J (2008) Mitochondrial DNA diversity and taxa delineation in the land snails of the *Iberus qualterianus* (Pulmonata, Helicidae) complex. *Zool J Linn Soc* 154: 722–737.
107. Elejalde MA, Munoz B, Arrebola JR, Gomez-Moliner BJ (2005) Phylogenetic relationships of *Iberus qualterianus* and *I. alonensis* (Gastropoda : Helicidae) based on partial mitochondrial 16S rRNA and COI gene sequences. *J Molluscan Stud* 71: 349–355.
108. Fehér Z, Szabó K, Bozsó M, Péntes Z (2009) Phylogeny and phylogeography of the *Lozekia-Kovacsia* species group (Gastropoda: Hygromiidae). *J Zool Syst Evol Res* 47: 306–314.
109. Van Riel P, Jordaens K, Van Houtte N, Martins AMF, Verhagen R, et al. (2005) Molecular systematics of the endemic *Leptaxini* (Gastropoda: Pulmonata) on the Azores islands. *Mol Phylogenet Evol* 37: 132–143.
110. Fiorentino V, Salomone N, Manganeli G, Giusti F (2008) Phylogeography and morphological variability in land snails: the Sicilian *Marmorana* (Pulmonata, Helicidae). *Biol J Linn Soc* 94: 809–823.
111. Moussalli A, Herbert DG, Stuart-Fox D (2009) A phylogeny of the cannibal snails of southern Africa, genus *Natalina* sensu lato (Pulmonata: Rhytididae): Assessing concordance between morphology and molecular data. *Mol Phylogenet Evol* 52: 167–182.
112. Weaver KF, Anderson T, Guralnick R (2006) Combining phylogenetic and ecological niche modeling approaches to determine distribution and historical biogeography of Black Hills mountain snails (Orcohelicidae). *Divers Distrib* 12: 756–766.
113. Lee T, Burch JB, Coote T, Fontaine B, Gargominy O, et al. (2007) Prehistoric inter-archipelago trading of Polynesian tree snails leaves a conservation legacy. *Proc R Soc Biol Sci Ser B* 274: 2907–2914.
114. Lee TW, Burch JB, Jung YH, Coote T, Pearce-Kelly P, et al. (2007) Tahitian tree snail mitochondrial clades survived recent mass extirpation. *Curr Biol* 17: R502–R503.
115. Rundell RJ, Holland BS, Cowie RH (2004) Molecular phylogeny and biogeography of the endemic Hawaiian Succineidae (Gastropoda : Pulmonata). *Mol Phylogenet Evol* 31: 246–255.
116. Ponder WF, Colgan DJ, Gleeson DM, Sherley GH (2003) Relationships of *Placostylus* from Lord Howe Island: an investigation using the mitochondrial cytochrome c oxidase 1 gene. *Molluscan Res* 23: 159–178.
117. Wilke T, Duncan N (2004) Phylogeographical patterns in the American Pacific Northwest: lessons from the arionid slug *Prophysaon coeruleum*. *Mol Ecol* 13: 2303–2315.
118. von Proschwitz T, Schander C, Jueg U, Thorkildsen S (2009) Morphology, ecology and DNA-barcoding distinguish *Pupilla pratensis* (Clessin, 1871) from *Pupilla muscorum* (Linnaeus, 1758) (Pulmonata: Pupillidae). *J Molluscan Stud* 75: 315–322.
119. Kameda YC, Kawakita A, Kato M (2007) Cryptic genetic divergence and associated morphological differentiation in the arboreal land snail *Satsuma (Luchuhadra) largillierti* (Camaenidae) endemic to the Ryukyu Archipelago, Japan. *Mol Phylogenet Evol* 45: 519–533.
120. Ketmaier V, Giusti F, Caccone A (2006) Molecular phylogeny and historical biogeography of the land snail genus *Solatopupa* (Pulmonata) in the peri-Tyrrhenian area. *Mol Phylogenet Evol* 39: 439–451.
121. Holland BS, Cowie RH (2007) A geographic mosaic of passive dispersal: population structure in the endemic Hawaiian amber snail *Succinea caduca* (Mighels, 1845). *Mol Ecol* 16: 2422–2435.
122. Depraz A, Cordellier M, Hausser J, Pfenninger M (2008) Postglacial recolonization at a snail's pace (*Trochulus villosus*): confronting competing refugia hypotheses using model selection. *Mol Ecol* 17: 2449–2462.
123. Nekola JC, Coles BF, Berghorsson U (2009) Evolutionary pattern and process within the *Vertigo gouldii* (Mollusca: Pulmonata, Pupillidae) group of minute North American land snails. *Mol Phylogenet Evol* 53: 1010–1024.
124. Efford M, Howitt R, Gleeson D (2002) Phylogenetic relationships of *Waimia* (Mollusca : Pulmonata) - biogeography and conservation implications. *J R Soc N Z* 32: 445–456.
125. Sauer J, Hausdorf B (2009) Sexual selection is involved in speciation in a land snail radiation on Crete. *Evolution* 63: 2535–2546.
126. Huggall A, Moritz C, Moussalli A, Stanisc J (2002) Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier 1875). *Proc Natl Acad Sci U S A* 99: 6112–6117.
127. Ueshima R, Asami T (2003) Single-gene speciation by left-right reversal - A land-snail species of polyphyletic origin results from chirality constraints on mating. *Nature* 425: 679–679.
128. Yamazaki N, Ueshima R, Terrett JA, Yokobori S, Kaifu M, et al. (1997) Evolution of pulmonate gastropod mitochondrial genomes: Comparisons of gene organizations of *Euhadra*, *Cepaea* and *Albinaria* and implications of unusual tRNA secondary structures. *Genetics* 145: 749–758.
129. Douris V, Giokas S, Thomaz D, Lecanidou R, Rodakis GC (2007) Inference of evolutionary patterns of the land snail *Albinaria* in the Aegean archipelago: Is vicariance enough? *Mol Phylogenet Evol* 44: 1224–1236.
130. Quinteiro J, Rodriguez-Castro J, Castillejo J, Iglesias-Pineiro J, Rey-Mendez M (2005) Phylogeny of slug species of the genus *Arion*: evidence of Iberian endemics and of the existence of relict species in Pyrenean refuges. *J Zool Syst Evol Res* 43: 139–148.