Check for updates

RESEARCH ARTICLE

REVISED Genomic evidence for the parallel regression of melatonin synthesis and signaling pathways in placental mammals

[version 2; peer review: 2 approved]

Christopher A. Emerling[1-3], Mark S. Springer[4], John Gatesy[5], Zachary Jones[1], Deana Hamilton[1], David Xia-Zhu[1], Matt Collin[4], Frédéric Delsuc[2]

[1]Museum of Vertebrate Zoology, University of California, Berkeley, Berkeley, CA, 94720, USA
[2]Institut des Sciences de l'Evolution de Montpellier (ISEM), CNRS, IRD, EPHE, Université de Montpellier, Montpellier, France
[3]Biology Department, Reedley College, Reedley, CA, 93654, USA
[4]Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, Riverside, CA, 92521, USA
[5]Division of Vertebrate Zoology, American Museum of Natural History, New York, NY, 10024, USA

v2 **First published:** 01 Jul 2021, **1**:75
https://doi.org/10.12688/openreseurope.13795.1
Latest published: 13 Dec 2021, **1**:75
https://doi.org/10.12688/openreseurope.13795.2

Abstract

Background: The study of regressive evolution has yielded a wealth of examples where the underlying genes bear molecular signatures of trait degradation, such as pseudogenization or deletion. Typically, it appears that such disrupted genes are limited to the function of the regressed trait, whereas pleiotropic genes tend to be maintained by natural selection to support their myriad purposes. One such set of pleiotropic genes is involved in the synthesis (*AANAT*, *ASMT*) and signaling (*MTNR1A*, *MTNR1B*) of melatonin, a hormone secreted by the vertebrate pineal gland. Melatonin provides a signal of environmental darkness, thereby influencing the circadian and circannual rhythmicity of numerous physiological traits. Therefore, the complete loss of a pineal gland and the underlying melatonin pathway genes seems likely to be maladaptive, unless compensated by extrapineal sources of melatonin.

Methods: We examined *AANAT*, *ASMT*, *MTNR1A* and *MTNR1B* in 123 vertebrate species, including pineal-less placental mammals and crocodylians. We searched for inactivating mutations and modelled selective pressures (dN/dS) to test whether the genes remain functionally intact.

Results: We report that crocodylians retain intact melatonin genes and express *AANAT* and *ASMT* in their eyes, whereas all four genes have been repeatedly inactivated in the pineal-less xenarthrans, pangolins, sirenians, and whales. Furthermore, colugos have lost these genes, and several lineages of subterranean mammals have partial melatonin pathway dysfunction. These results are supported by the presence of shared inactivating mutations across clades and

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 13 Dec 2021	✓ view	✓ view
version 1 01 Jul 2021	? view	? view

1. **Mareike C. Janiak**, University of Salford, Salford, UK

2. **Ryan K Schott**, York University, York, Canada

Any reports and responses or comments on the article can be found at the end of the article.

analyses of selection pressure based on the ratio of non-synonymous to synonymous substitutions (dN/dS), suggesting extended periods of relaxed selection on these genes.
Conclusions: The losses of melatonin synthesis and signaling date to tens of millions of years ago in several lineages of placental mammals, raising questions about the evolutionary resilience of pleiotropic genes, and the causes and consequences of losing melatonin pathways in these species.

Keywords
Melatonin, Pseudogene, Xenarthra, Pholidota, Dermoptera, Sirenia



This article is included in the Horizon 2020 gateway.



This article is included in the Marie-Sklodowska-Curie Actions (MSCA) gateway.



This article is included in the Genetics and Genomics gateway.



This article is included in the European Research Council (ERC) gateway.



This article is included in the Evolution and Ecology gateway.



This article is included in the Evolutionary Biology collection.

Corresponding author: Christopher A. Emerling (christopher.emerling@reedleycollege.edu)

Author roles: Emerling CA: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Springer MS**: Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Writing – Review & Editing; **Gatesy J**: Data Curation, Funding Acquisition, Investigation, Writing – Review & Editing; **Jones Z**: Formal Analysis, Investigation, Writing – Review & Editing; **Hamilton D**: Formal Analysis, Investigation, Writing – Review & Editing; **Xia-Zhu D**: Formal Analysis, Investigation, Writing – Review & Editing; **Collin M**: Data Curation, Investigation; **Delsuc F**: Data Curation, Funding Acquisition, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the European Research Council (ERC) grant agreement No [683257]. This research was also supported by a National Science Foundation (NSF) Postdoctoral Research Fellowship in Biology (award no. 1523943; C.A.E.), an NSF Postdoctoral Fellow Research Opportunities in Europe award (C.A.E.), the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. PCOFUND-GA-2013-609102, through the PRESTIGE programme coordinated by Campus France (C.A.E.) and an NSF DEB- 1457735 grant to J.G and M.S.S.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Emerling CA *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 work is properly cited.

How to cite this article: Emerling CA, Springer MS, Gatesy J *et al.* **Genomic evidence for the parallel regression of melatonin synthesis and signaling pathways in placental mammals [version 2; peer review: 2 approved]** Open Research Europe 2021, **1**:75 https://doi.org/10.12688/openreseurope.13795.2

First published: 01 Jul 2021, **1**:75 https://doi.org/10.12688/openreseurope.13795.1

REVISED **Amendments from Version 1**

This manuscript has been revised in a few important ways: 1. we elaborated on the methodology provided for the evolutionary analyses (specifically dN/dS analyses), 2. we performed new analyses on historical evolutionary pressures of crocodylian melatonin genes, 3. we provided additional supporting information validating the pseudogene status of nearly every putatively inactivated melatonin pathway gene in our focal taxa. Beyond this, we modified text and added supplementary information to provide further clarity to our study. None of the results changed in any significant way from what was reported in the first published manuscript. If anything, our conclusions have been strengthened by the addition of new data and analyses.

Any further responses from the reviewers can be found at the end of the article

Plain language summary

Evolution is typically thought to occur by making an organism more complex, such as through the addition of new traits. However, evolution can also proceed through the loss or degeneration of characteristics. The reduction of eyes in animals living underground and the loss of limbs in snakes are examples of such regressive evolution. When organisms evolve regressively, the genes underlying such traits often become broken (pseudogenes) or deleted from their genome altogether. However, genes are not typically thought to be lost when they perform many functions (pleiotropy). For instance, four genes are involved in the production (*AANAT*, *ASMT*) and detection (*MTNR1A*, *MTNR1B*) of melatonin, a hormone that signals an animal's organ systems that it is dark outside (e.g., during the night). Melatonin genes seem unlikely to become broken or lost given that this hormone is responsible for signalling many parts of the body to perform functions specific to nighttime. The loss of such genes would likely have a negative, cascading effect on many body functions; therefore, natural selection would presumably retain functional versions of these genes. Surprisingly, however, certain vertebrates are reported to lack the organ responsible for secreting melatonin (pineal gland). We examined the melatonin genes in these vertebrates to determine if they are broken, missing and/or show evidence of degradation over time. We discovered that while pineal-less crocodylians retain all four melatonin genes, other species reported to lack a pineal gland (anteaters, sloths, armadillos, pangolins, dugong, manatee, whales) have broken and/or missing melatonin genes. Furthermore, colugos (flying lemurs) and some mammals that live underground in complete darkness show evidence of melatonin gene dysfunction. Together, these results indicate that these mammals lost their ability to produce and detect melatonin tens of millions of years ago, raising questions as to how they have adapted to a life without this hormone.

Introduction

Evidence for the molecular basis of regressive evolution, or vestigialization, has become abundant following increases in the availability of whole genome assemblies (Albalat & Cañestro, 2016). Such molecular regression typically manifests as the accumulation of unitary pseudogenes and whole gene deletions, although an increasing number of studies are finding that it can also result in or arise from the erosion of gene promoters and enhancers. Among vertebrates, examples include mutations associated with the loss of teeth in exchange for beaks and baleen (Deméré *et al.*, 2008; Meredith *et al.*, 2014), reduction in visual capabilities as species eschew life aboveground in favor of a subterranean existence (Emerling & Springer, 2014; Fang *et al.*, 2014; Partha *et al.*, 2017), and the loss or reduction of limbs while adapting to new locomotory strategies (Berger *et al.*, 2018; Emerling, 2017a; Sackton *et al.*, 2019; Vonk *et al.*, 2013). A unifying theme is that the loci underlying the loss of these traits typically appear to be restricted to a single or limited function: tooth genes specifically associated with enamel and dentin development become inactivated in toothless vertebrates (Meredith *et al.*, 2014), whereas tooth genes with bone-related or other functions remain intact (Deméré *et al.*, 2008; Springer *et al.*, 2016); genes encoding light-sensitive opsins used in bright light conditions are lost as taxa adapt to life underground (David-Gray *et al.*, 2002; Emerling & Springer, 2014), but genes otherwise necessary for eye formation remain intact (Emerling, 2018; Partha *et al.*, 2017); and while a gene encoding a claw-specific keratin in squamates is pseudogenized in snakes (Emerling, 2017a), multipurpose *Hox* genes associated with limb-development typically remain conserved (Vonk *et al.*, 2013). Such observations should not be surprising given that the pleiotropic nature of many genes necessitates their retention, even if one or more associated traits are lost. In such pleiotropic genes, the loss of regulatory non-coding elements associated with specific functions appears more likely to occur than outright disruption of gene function (Berger *et al.*, 2018; Kvon *et al.*, 2016; Partha *et al.*, 2017). However, reports of the absence of the pineal gland in several lineages of vertebrates (Oksche, 1965) challenge the assumption that pleiotropic effects lead to gene conservation.

The pineal gland is an endocrine organ located within the diencephalon of vertebrates, whose primary, and perhaps only, function is to secrete melatonin. Melatonin functions as a potent antioxidant and can also act as a hormone that signals environmental darkness (Hardeland & Poeggeler, 2003; Zhao *et al.*, 2019). In vertebrates, melatonin derives from serotonin, which is converted into *N*-Acetylserotonin by aralkylamine *N*-acetyltransferase (AANAT) and then modified into melatonin by *N*-Acetylserotonin O-methyltransferase (ASMT/HIOMT). Melatonin synthesis in the pineal gland is under the control of the circadian master clock (suprachiasmatic nucleus), following a pattern of high production in darkness and low production in light. It is released into the bloodstream and arrives at target tissues to activate downstream pathways via the G protein-coupled melatonin receptors type 1A (MTNR1A/MT1) and 1B (MTNR1B/MT2) (Cipolla-Neto *et al.*, 2014). These melatonin receptors are expressed widely, including in the suprachiasmatic nucleus, thalamus, cerebral cortex, retina, kidneys, adrenal glands, reproductive organs, arteries, immune cells, liver, pancreas, skin and bone, indicating broad signaling from this hormone (Cipolla-Neto & Do Amaral, 2018;

Sapède & Cau, 2013). Indeed, by providing a signal of darkness, melatonin helps to maintain circadian and circannual rhythms, influencing energy metabolism, seasonal reproduction, migration behavior, blood pressure, immune system functioning, among other processes (Cipolla-Neto *et al*., 2014; Nishiwaki-Ohkawa & Yoshimura, 2016; Sapède & Cau, 2013).

Despite such seemingly indispensable effects, and the experience of a light/dark cycle by nearly all vertebrates, a number of vertebrates are reported to lack a pineal gland: sloths, armadillos, and anteaters (Xenarthra), pangolins (Pholidota), certain whales (Cetacea), the dugong (*Dugong dugon*; Sirenia) and crocodylians (McFarland *et al*., 1969; Oelschläger *et al*., 2008; Oksche, 1965; Panin *et al*., 2012; Phillips *et al*., 1986; Ralph *et al*., 1985). If accurate, the phenotypic consequences of being pineal-less would likely be widespread, as shown by how pinealectomies can impact clock gene expression (Coelho *et al*., 2015; De Farias *et al*., 2015), insulin function (Lima *et al*., 1998), serum leptin levels (Canpolat *et al*., 2001), dopamine levels (Khaldy *et al*., 2002), spinal development (Fjelldal *et al*., 2004), gonadal function (Baburski *et al*., 2015; Haldar & Ghosh, 1990), and immune function (Del Gobbo *et al*., 1989). However, despite such anatomical observations, scientists have reported circulating serum melatonin in the nine-banded armadillo (*Dasypus novemcinctus*), bottlenose dolphin (*Tursiops truncatus*), American alligator (*Alligator mississippiensis*), and freshwater crocodile (*Crocodylus johnstoni*) (Firth *et al*., 2010; Funasaka *et al*., 2011; Harlow *et al*., 1981; Panin *et al*., 2012; Roth *et al*., 1980), even reporting rhythmic secretions in some species. Furthermore, bottlenose dolphins reportedly show ASMT presence in several tissues (Panin *et al*., 2012), and exogenous melatonin evidently influences activity patterns and body temperature in the nine-banded armadillo (Harlow *et al*., 1982). However, a review on melatonin binding sites in the pars tuberalis, suprachiasmatic nucleus, hippocampus and other brain sites found low to no melatonin binding in the nine-banded armadillo, in contrast to 14 other mammals (Bittman, 1993).

Perhaps reconciling these apparently conflicting data, research has revealed that multiple organs in vertebrates can synthesize melatonin, such as the retina and lens of the eye, thymus and bone marrow (Acuña-Castroviejo *et al*., 2014; Cogburn *et al*., 1987; Gern *et al*., 1978a; Reppert & Sagar, 1983; Stefulj *et al*., 2001; Underwood *et al*., 1984). Indeed, there is evidence that extra-pineal melatonin can enter blood circulation (Foà & Menaker, 1988; Gern *et al*., 1978b; Gern & Norris, 1979; Kennaway *et al*., 1977; Lynch *et al*., 1975; Owens & Gern, 1985; Reppert & Sagar, 1983; Underwood *et al*., 1984), suggesting that extra-pineal organs may ultimately be the source of serum melatonin in pineal-less vertebrates. Furthermore, melatonin synthesis has also been reported in mitochondria (Suofu *et al*., 2017), providing another potential source of extra-pineal melatonin.

Given the ubiquity and myriad effects of this hormone, we hypothesized that melatonin synthesis has been maintained even in pineal-less vertebrates, but may rely on extra-pineal sources to perform the same functions. We set out to determine if the genes necessary for melatonin synthesis (*AANAT*, *ASMT*) and signaling (*MTNR1A, MTNR1B*) are functionally intact in several clades of apparently pineal-less vertebrates. We found evidence that in contrast to crocodylians, which maintain intact melatonin pathway genes, numerous placental mammal lineages show evidence of melatonin synthesis and/or signaling disruption. We inferred that several of these events occurred tens of millions of years ago, raising questions about the evolutionary resilience of pleiotropic systems.

Materials and methods

Gene dataset assembly

We obtained *AANAT*, *ASMT*, *MTNR1A* and *MTNR1B* gene sequences for 110 species of placental mammals, including 13 xenarthrans, three pangolins, three sirenians and 25 cetaceans, as well as 13 crocodylians (Underlying data, Supplementary Tables S1–S5 [Emerling *et al*., 2021]). Gene assembly was accomplished using a combination of whole genome assemblies and mapping short reads of sequenced genomes onto reference sequences using published and novel sequences.

We began by downloading human and chicken reference mRNA sequences for all four melatonin genes from GenBank (Underlying data, Supplementary Tables S2–S5 [Emerling *et al*., 2021]). The protein-coding sequence of each mRNA was BLASTed (megablast) against human and chicken genome assemblies in NCBI's (National Center for Biotechnology Information) WGS (Whole Genome Shotgun) database. For each gene, we downloaded a contiguous sequence that included all of the mRNA coding exons and flanking sequence, imported the sequences into Geneious v. 2019.2.3 (Kearse *et al*., 2012), and aligned each mRNA to its corresponding WGS contig hits using MUSCLE v. 3.5 (Edgar, 2004). We then used the human and chicken assembly-derived sequences as the references for obtaining orthologs in mammals and crocodylians, respectively.

For whole genome assemblies, we BLASTed our reference sequences against assemblies in external and imported databases (Underlying data, Supplementary Table S1 [Emerling *et al*., 2021]) using intermediate sensitivity settings (e.g., discontiguous megablast on NCBI). When obtaining *AANAT* sequences, the whole reference sequence was used as a probe in BLAST searches and short read mapping. *ASMT* sequences were obtained using a mixture of single exons plus flanking DNA and whole gene reference sequences due to the relatively large size of the gene and small exons. In the case of some *ASMT* sequences, we queried NCBI's annotated scaffolds directly (RefSeq). *MTNR1A* and *MTNR1B* both have two coding exons separated by a large intron, so to avoid incompatible homology issues when attempting to align the introns, we typically designed separate probes with flanking sequences for the two coding exons. If we ever failed to recover sequences for a species, we used mRNA sequences and annotated gene predictions, plus newly-assembled sequences, as reference sequences for BLAST and mapping, especially from close relatives. Sequences derived from genome assemblies that contain

long stretches of Ns can cause issues with alignments, so we trimmed any such instances to ten Ns.

Short reads were obtained from both published (NCBI, Sequence Read Archive) and novel sources. To generate novel sequences for multiple crocodylians (*Caiman latirostris*, *Crocodylus niloticus*, *Mecistops* sp., *Melanosuchus niger*, *Osteolaemus tetraspis*, *Paleosuchus palpebrosus*, *P. trigonatus*, *Tomistoma schlegelii*), high quality DNA was extracted from tissue using the Qiagen DNeasy kit and quantified through Qubit. In preparation for library construction, 200ng of DNA were sheared with a Covaris S220 for an average size of 500 bp and used as input for the Illumina Neoprep automated library construction instrument. Constructed libraries were pooled and sent to the New York Genome Center for sequencing on HiSeq X system paired-end 150 bp reads.

We imported short reads into Geneious and mapped them to reference sequences (exon + flanking) that were taxonomically close (i.e., same genus, family or order) using intermediate sensitivity settings (medium-low sensitivity). For each set of probes, we performed an initial mapping run to expedite sequence capture, followed by a remapping of the captured reads with the "Fine Tuning" option set to iterate up to five times. Mapped short reads were examined by eye and trimmed of nonhomologous sequences likely derived from adapters or sequencing errors.

For dN/dS analyses in crocodylians, we obtained sauropsid outgroup comparison sequences derived from predicted gene models only (Underlying data, Supplementary Tables S2–S5 [Emerling *et al.*, 2021]), not constructing any such outgroup sequences from assemblies. We BLASTed (discontiguous megablast) each *Alligator mississippiensis* ortholog against NCBI's nucleotide collection, restricting searches to Squamata, Testudinata and Aves, respectively.

Upon obtaining the mammalian and sauropsid gene sequences, each sequence was imported into Geneious and aligned to its respective probe (i.e., human or chicken whole gene sequence) using MUSCLE. When working with different clades, alignments for each species in that clade were performed successively to provide anchoring and improve subsequent alignments, followed by manual adjustment (see clades in the underlying data, Supplementary Tables S1–S5 [Emerling *et al.*, 2021]). After creating each clade-specific alignment (Underlying data, Supplementary Dataset S1 [Emerling *et al.*, 2021]), we examined the sequences for frameshift indels, splice donor mutations, and splice acceptor mutations. After excising the introns, the coding sequence was translated to search for start codon mutations, premature stop codons, and ancestral stop mutations (Underlying data, Supplementary Tables S2–S5 [Emerling *et al.*, 2021]).

Finally, following the completion of our analyses, a whole genome assembly (WGS) was released for the sirenian Steller's sea cow (*Hydrodamalis gigas*). The high similarity to other sirenian gene sequences (pairwise comparison with *Dugong dugon*: *AANAT*: 98%; *ASMT*: 97.4%; *MTNR1A*: 96%; *MTNR1B*: 96.6%) allowed us to positively identify these genes and characterize their functionality, but we did not include them in the evolutionary analyses.

Evolutionary analyses

Following examination of each clade-specific alignment, global alignments for each gene were assembled to perform phylogenetic (RAxML) and selection pressure (PAML) analyses, respectively (Underlying data, Supplementary Datasets S2–S8 [Emerling *et al.*, 2021]). All stop and incomplete codons were replaced with gaps, as were any codons in individual sequences that were difficult to align with confidence. RAxML analyses were performed with v. 8.2.12 (Stamatakis, 2014) on CIPRES (RAxML-HPC2 on XSEDE) (Miller *et al.*, 2010), using the default parameters (GTR-CAT), and executing 500 bootstrap replications.

We performed dN/dS ratio analyses using codeml in PAML ver. 4.8 (Yang, 2007) to test whether genes that appear to be pseudogenized show evidence of undergoing relaxed selection. The mammalian tree topology for these analyses (Underlying data, Supplementary Dataset S9 [Emerling *et al.*, 2021]) was largely derived from a single phylogeny (Emerling *et al.*, 2015), with cetacean, carnivoran, xenarthran and talpid relationships obtained from additional sources (Allio *et al.*, 2021; Gibb *et al.*, 2016; He *et al.*, 2017; McGowen *et al.*, 2020). In some cases, relationships for certain mammals were not resolved in the reference trees, but confamilials and/or congeners were present, allowing us to confidently place such taxa in the phylogeny. For example, our primary reference tree does not include *Desmodus rotundus*, but other phyllostomid species are present. Therefore, *D. rotundus* was positioned in the phyllostomid portion of the PAML tree topology. After setting the topology, we executed one-ratio branch models with one of three codon frequency models (1/61 each, F1X4, F3X4) and used the Akaike information criterion to select the best-fitting codon frequency model for each gene (Underlying data, Supplementary Table S6 [Emerling *et al.*, 2021]).

To test whether inferred pseudogenes are under relaxed selection, we employed branch tests on a series of nested models relative to a 'master' model to estimate whether ω is elevated on relevant branches. To construct the 'master' models, we used an approach similar to that employed by Meredith *et al.* (2009), categorizing branches as either functional, pseudogenic, mixed/transitional or pre-mutation. 'Functional' branches are those for which there is no evidence for gene inactivation, and there is no expectation that the gene is likely to become a pseudogene. All 'functional' branches were grouped together for a single ω, were expected to be consistent with purifying selection ($\omega < 1$) and therefore were treated as the background with which to compare the remaining three branch categories. 'Pseudogenic' branches are those which post-date a branch on which a gene is inferred to have become a pseudogene. For example, if we mapped an inactivating mutation to the stem cetacean branch, all of the descendant branches within crown Cetacea were grouped together, and a single ω was estimated for this set of branches. In such cases, the dN/dS is expected

to approach or be under relaxed selection (ω = 1, higher than functional category). 'Mixed/transitional' branches are those on which a gene is inferred to have become a pseudogene, due to one or more inactivating mutations being mapped to that branch. It is therefore suspected to have a mixed history, transitioning from a functional gene to a pseudogene. Depending on how early or late this transition has occurred, the ω is expected to be highly similar to the functional category (more recent inactivation), very similar to the pseudogene category (very early inactivation) or in between these two extremes. Finally, the 'pre-mutation' branch category represents instances in which we found no positive evidence of pseudogenization on a branch, but external evidence suggests that relaxed selection may have occurred. Such pre-mutation branches could be designated due to multiple immediate descendant lineages having inactivating mutations and/or one or more other key genes being inactivated. As an example of the former, if a gene is inferred to have been pseudogenized within sloths, armadillos and anteaters separately, given that these all belong to the xenarthran clade, we would designate the branches prior to the mixed/transitional branches of these subclades as 'pre-mutation' branches. As an example of the latter, if a taxon had one or both melatonin synthesis genes (e.g., *Fukomys damarensis*) and/or both melatonin receptor genes (e.g., *Heterocephalus glaber*) inactivated, but no evidence of pseudogenization in the remaining genes, we would categorize these branches as 'pre-mutation'.

As we were not interested in average selection patterns across all branches within a particular category, except the background (functional) branches, we allowed each pre-mutation branch, each mixed/transitional branch and each set of pseudogene branches that post-dated an inactivation event to have their own ω estimates. This allowed us to test whether unique historical instances of putative gene inactivation were consistent with relaxed selection, rather than obscuring the signal by 'averaging' pre-mutation, mixed/transitional and pseudogene branches, respectively, across the tree.

After setting these branch categories, this 'master' model (Extended data, Supplementary Figures S1–S4 [Emerling *et al.*, 2021]), was calculated for each gene and used as the basis for which to compare subsequent nested models (Underlying data, Supplementary Tables S7–S10 [Emerling *et al.*, 2021])). Specifically, if higher than the background in the master model, we tested whether each set of pseudogene category branches, each mixed/transitional branch and each pre-mutation branch was distinguishable from the background (functional) ω and/or from the neutral ω value of 1 (expectation for pseudogene branches). In such cases, we used the master model for each gene that included all branch categories (Extended data, Supplementary Figures S1–S4 [Emerling *et al.*, 2021]), and would change only one branch of interest to be fixed as being part of the background or 1. We then compared each of the nested models to the master model using likelihood ratio tests to determine which models better fit the data for a branch of interest. In these instances, the models in which branch(es) were fixed as the background or 1 are considered the null model, with the master model being the alternative model, i.e., we tested whether a free estimate of ω for a foreground branch is statistically distinguishable from the background and/or 1. Finally, given the number of comparisons we accumulated, we ran Holm-Bonferroni corrections for multiple testing (Gaetano, 2018; Holm, 1979) (Underlying data, Supplementary Tables S7–10 [Emerling *et al.*, 2021]).

We performed separate dN/dS ratio analyses for crocodylian melatonin genes by comparing their pattern of molecular evolution to those of 19 other sauropsids (five squamates, five turtles, nine birds). The topology for our analyses (Underlying data, Supplementary Dataset S9 [Emerling *et al.*, 2021]) is a composite derived from separate phylogenies for crocodylians (Hekkala *et al.,* 2021), squamates (Zheng & Wiens, 2016), turtles (Pereira *et al.*, 2017) and birds (Prum *et al.*, 2015). As above, we first chose the codon frequency model for each gene by running one ratio models and determining the best fit using the Akaike information criterion (Underlying data, Supplementary Table S6 [Emerling *et al.*, 2021]). Then we ran a three-ratio model, estimating ω separately for crown Crocodylia, stem Crocodylia and the background, and compared its fit to a one ratio model using a likelihood ratio test (Underlying data, Supplementary Table S11 [Emerling *et al.*, 2021]).

Melatonin gene expression

In order to test for evidence of active transcription of the melatonin synthesis genes in a crocodylian, we analyzed published RNA sequencing data from 22 sequencing experiments on tissues in a juvenile *Alligator mississippiensis* (St John *et al.*, 2012). We BLASTed (megablast) the protein-coding regions of the *A. mississippiensis AANAT* and *ASMT* genes against short reads derived from mRNA sequencing in NCBI's Sequence Read Archive (Underlying data, Supplementary Table S12 [Emerling *et al.*, 2021]), imported the sequences into Geneious, and mapped them to the *A. mississippiensis* references using the low sensitivity setting.

Results

We found evidence that all xenarthrans, all pangolins, and all cetaceans have had both their melatonin synthesis (*AANAT*, *ASMT*) and melatonin signaling (*MTNR1A*, *MTNR1B*) genes disrupted via the accumulation of inactivating mutations and whole gene deletions (Figure 1 & Figure 2; Underlying data, Supplementary Tables S2–S5 [Emerling *et al.*, 2021]). Among sirenians, the West Indian manatee (*Trichechus manatus*) and dugong (*Dugong dugon*) have inactivated *AANAT*, *MTNR1A* and *MTNR1B*, with the former being heterozygous for an *ASMT* pseudogene (Underlying data, Supplementary Table S13 [Emerling *et al.*, 2021]) and the latter retaining an intact *ASMT*, whereas the Steller's sea cow (*Hydrodamalis gigas*) only has an inactivated *MTNR1B*. By contrast, all 13 crocodylians investigated retain intact melatonin genes (Figure 3; Underlying data, Supplementary Tables S2–S5 [Emerling *et al.*, 2021]), which show evidence of being under purifying selection (Underlying data, Supplementary Table S11 [Emerling *et al.*, 2021]).

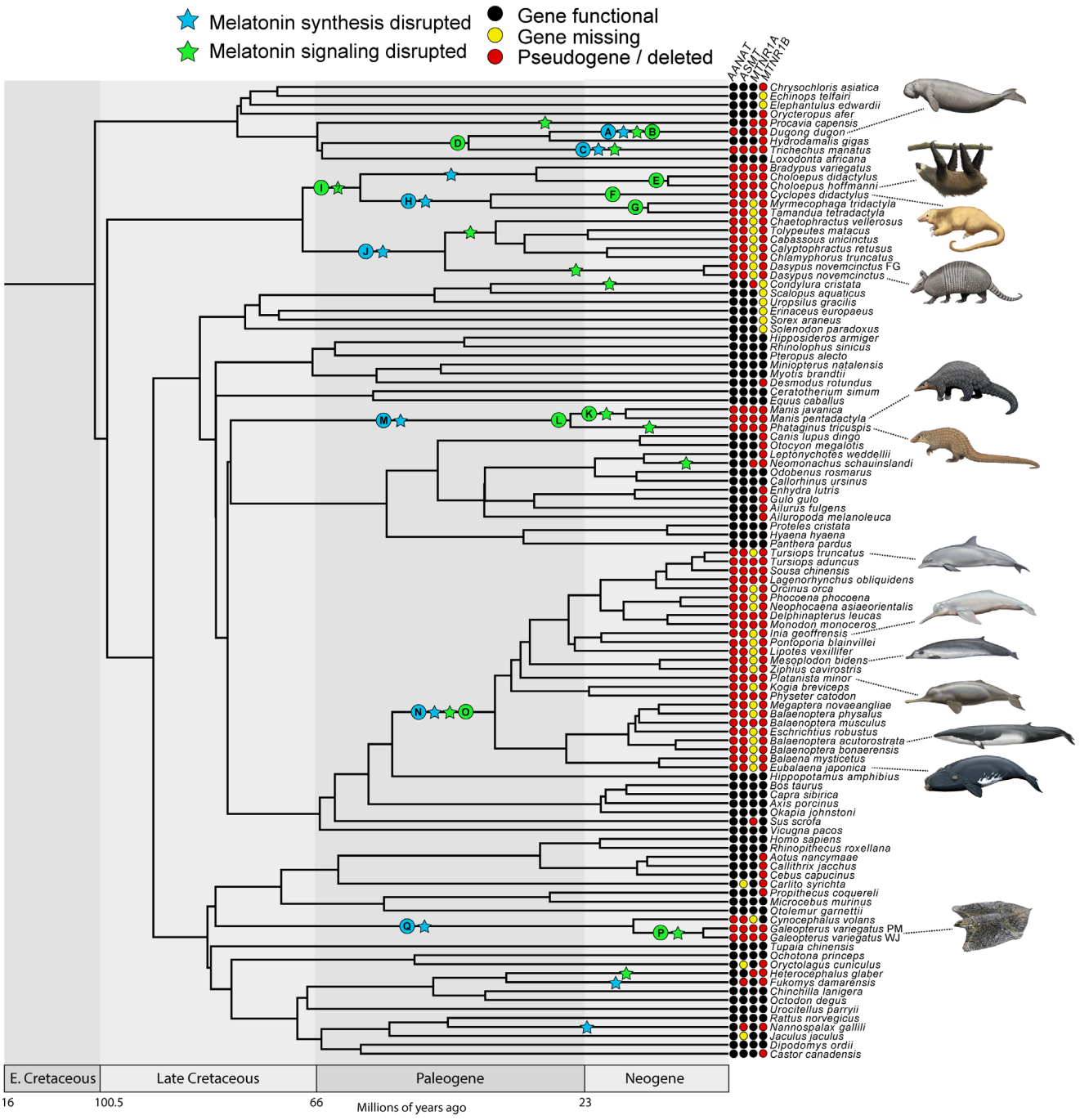


Figure 1. Timetree showing patterns of melatonin gene retention and loss across the placental mammals examined in this study. "Melatonin synthesis disrupted" indicates *AANAT* and/or *ASMT* is inferred to have been inactivated on the associated branch. "Melatonin signaling disrupted" indicates *MTNR1A* and *MTNR1B* are both inferred to have been inactivated on the associated branch, or one gene was lost on an earlier branch and the second was inactivated on the associated branch. Note that the stars are arbitrarily placed in the middle of branches and do not correspond to a precise timing for gene loss. Letters on stars and nodes correspond to letters in Figure 2. References for topology in Materials and Methods. Divergence dates in the figure derived from multiple references (Gibb *et al.*, 2016; Kumar *et al.*, 2017; McGowen *et al.*, 2020; Meredith *et al.*, 2011; Springer *et al.*, 2015). Paintings by Carl Buell, copyright John Gatesy.

Furthermore, we found that *AANAT* and *ASMT* are both expressed in the American alligator, particularly within the eye (Underlying data, Supplementary Table S12 [Emerling *et al.*, 2021]). We also found evidence for the complete disruption of both melatonin gene pathways (i.e., synthesis, signaling) in two dermopteran (colugo) species (*Galeopterus variegatus* Peninsular Malaysia, *G. variegatus* West Java; considered distinct cryptic species by Mason *et al.*, 2016), with a third (*Cynocephalus*

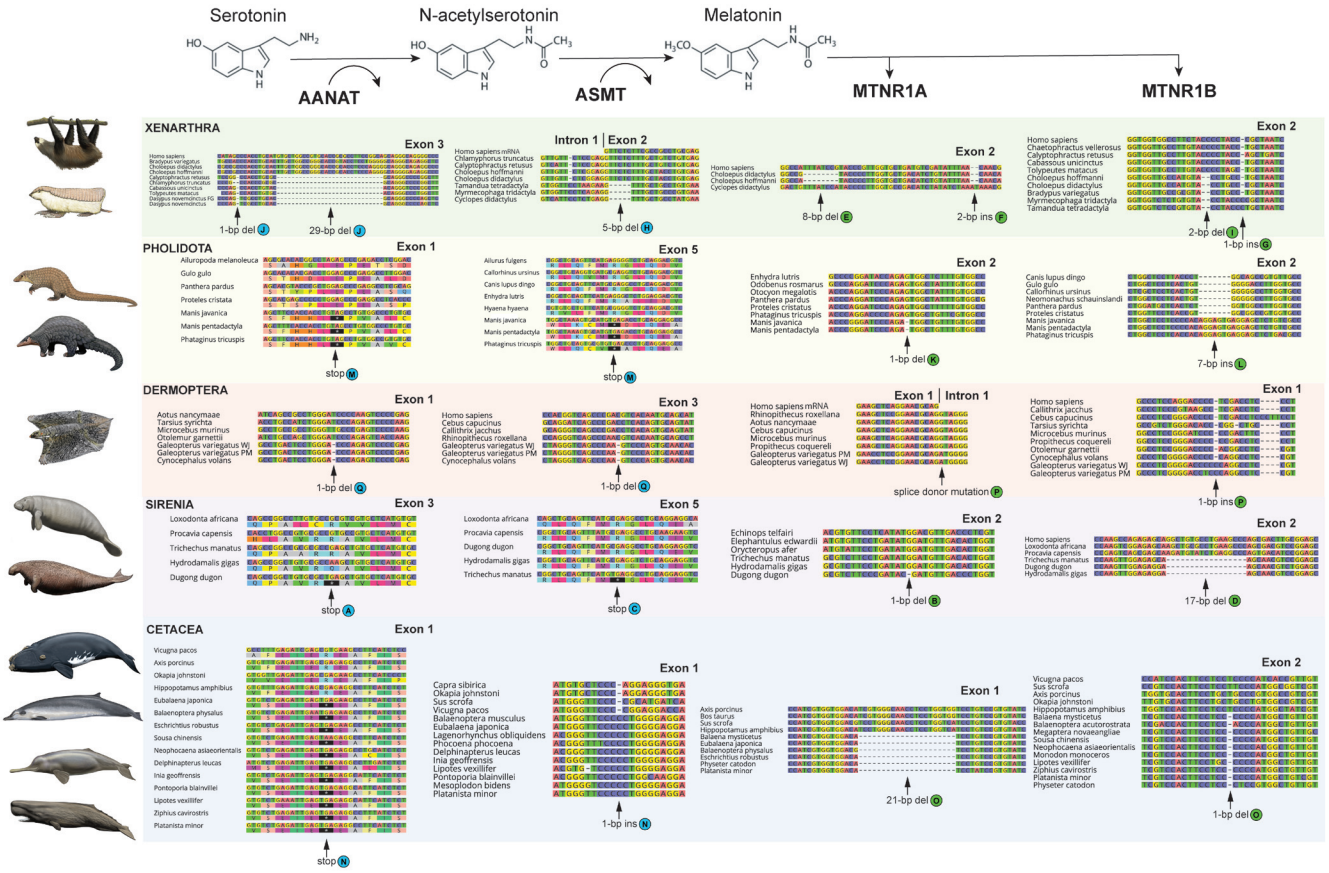


Figure 2. Sampling of inactivating mutations in melatonin pathway genes from five clades of placental mammals. Each column of DNA and protein sequence alignments corresponds to a bolded protein in the melatonin pathways towards the top of the figure. Letters after mutations correspond to letters in the timetree in Figure 1. Ins = insertion; del = deletion; stop = premature stop codon. Paintings by Carl Buell, copyright John Gatesy.

volans) possibly only retaining a functional *MTNR1B* (Figure 1 & Figure 2; Underlying data, Supplementary Tables S2–S5 [Emerling *et al.*, 2021]). The melatonin synthesis gene (*ASMT*) is also inactivated in three subterranean rodents (*Nannospalax galili*, *Spalax carmeli*, *Fukomys damarensis*), and both melatonin receptor genes appear to be inactivated in hyraxes (*Procavia capensis*, *Heterohyrax brucei*), a subterranean mole (*Condylura cristata*) and mole-rat (*Heterocephalus glaber*), and a seal (*Neomonachus schauinslandi*) (Underlying data, Supplementary Tables S3–S5, S13 [Emerling *et al.*, 2021]).

The pseudogene status of nearly all of these genes is supported by the positive identity of the genes as orthologs inferred by RAxML analyses (Extended data, Supplementary Figures S5–S7 [Emerling *et al.*, 2021]), the absence of functional paralogs (Extended data, Supplementary Figures S5–S7; Underlying data, Supplementary Tables S2–S5 [Emerling *et al.*, 2021]), presence of shared inactivating mutations (Figure 2; Underlying data, Supplementary Tables S2–S5, S13; Extended data, Supplementary Figures S8–S12 [Emerling *et al.*, 2021]), the mapping of short reads (Underlying data, Supplementary Table S13; Extended data, Supplementary Figures S9–S14 [Emerling *et al.*, 2021]) and dN/dS ratio estimates suggestive of extensive relaxed selection (Extended data, Supplementary Figures S15–S18; Underlying data, Supplementary Tables S7–S10 [Emerling *et al.*, 2021]). Specifics are discussed below.

Crocodylia

We were able to obtain whole or partial genes for all 13 crocodylians we examined, and none of the recovered sequences presented evidence of gene inactivation. Furthermore, dN/dS ratio analyses strongly suggest that all four melatonin genes are under purifying selection within crown Crocodylia (*AANAT*: background $\omega = 0.1$, crown Crocodylia $\omega = 0.1$; *ASMT*: background $\omega = 0.2$, crown Crocodylia $\omega = 0.15$; *MTRN1A*: background $\omega = 0.09$, crown Crocodylia $\omega = 0.12$; *MTNR1B*: background $\omega = 0.11$, crown Crocodylia $\omega = 0.13$; Underlying data, Supplementary Table S11 [Emerling *et al.*, 2021]). Out of 22 published RNA sequencing experiments on a juvenile *Alligator mississippiensis* (Underlying data, Supplementary Table S12 [Emerling *et al.*, 2021]), five yielded reads that mapped to *AANAT* and/or *ASMT*: eye (both), ovary (*ASMT*), pooled

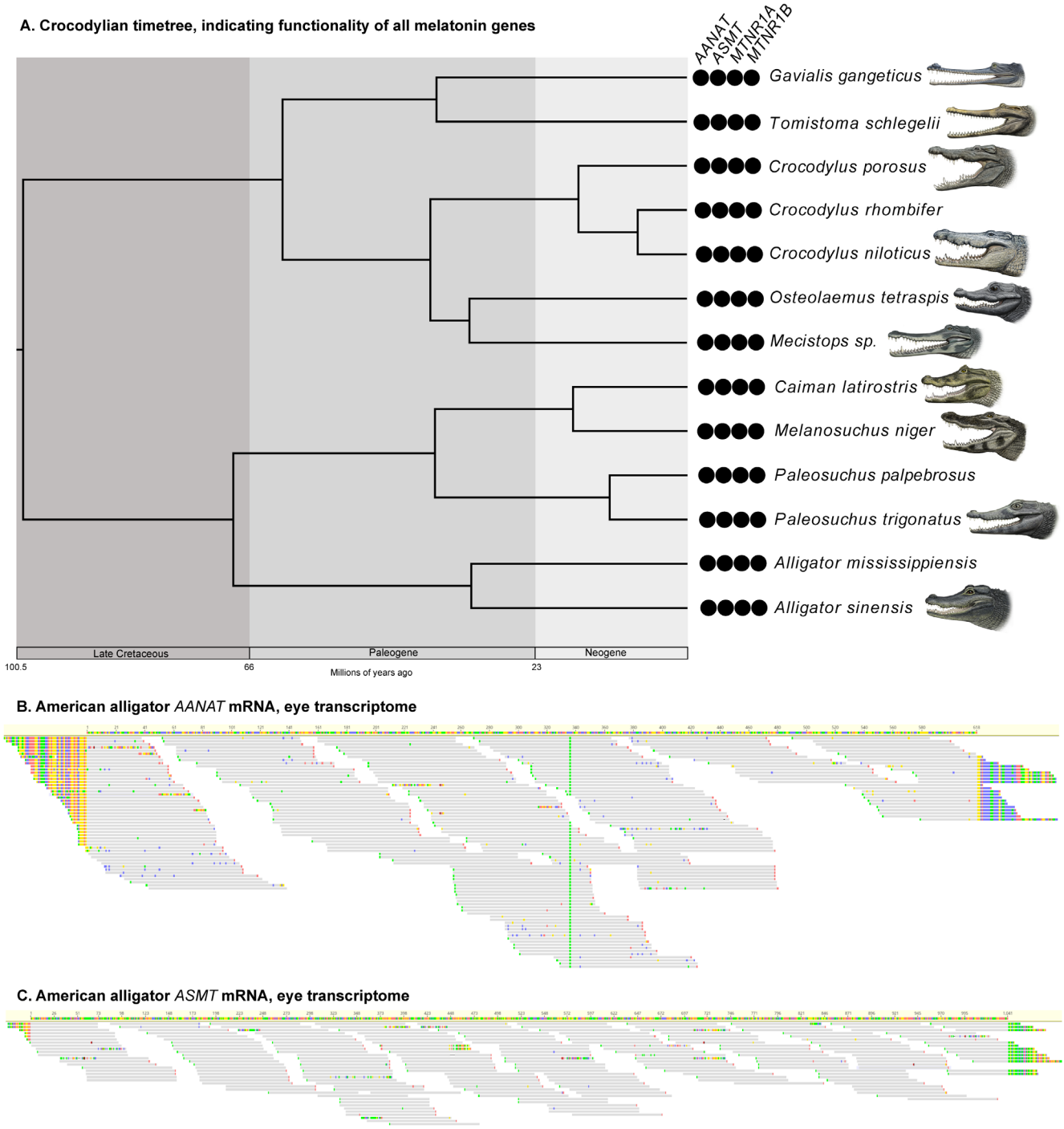


Figure 3. Melatonin genes in crocodylians. A. Timetree depicting relationships and divergence times of crocodylians (Hekkala *et al.*, 2021) examined in this study. The black circles indicate retention of functional orthologs of the genes. Paintings by Carl Buell, copyright John Gatesy. **B**. mRNA short read from an American alligator eye mapped to a reference *AANAT*. **C**. mRNA short reads from an American alligator eye mapped to a reference *ASMT*.

(intracoelomic fat body, scute muscle, trachea, cloacal gland, gastralia; both), stomach (*AANAT*) and thalamus (*ASMT*). The ovary, stomach and thalamus yielded only two reads each, the pooled experiment yielded 27 for *AANAT* and 22 for *ASMT*, and the eye had 251 for *AANAT* and 207 for *ASMT*. The mapped reads for the eye experiment encompassed the entire genes

(Figure 3), strongly suggesting that the melatonin synthesis genes are expressed in one or more ocular tissues.

Xenarthra

For *AANAT*, sloths (Folivora) share a mutated start codon (ACA), splice donor mutation (GT to GG) in intron 1, 1-bp deletion in exon 2, and two frameshift indels in exon 3. All anteaters (Vermilingua) possess a start codon mutation (GTG or TTG), with six additional inactivating mutations shared between *Myrmecophaga* and *Tamandua*. Armadillos (Cingulata) share a 1-bp deletion in exon 1, and four frameshift indels in exon 3 (Figure 2). In *ASMT*, there is a putative 1-bp insertion in exon 6 shared by two-toed sloths (*Choloepus* spp.), two chlamyphorid armadillos (*Calyptophractus*, *Chlamyphorus*), and a dasypodid armadillo (*Dasypus novemcinctus* French Guiana). This region is missing in all other xenarthran sequences (i.e., deleted or not assembled), and the alignment is admittedly ambiguous. Furthermore, in all xenarthrans we examined, exon 5 appears to be absent and the ancestral stop codon (i.e., the stop codon common to all other species examined) is mutated, although it differs among sloths (TAT), anteaters (TGT or TGG), and armadillos (CGT or TGT). Beyond these, shared inactivating mutations are present among the anteaters (5-bp deletion, exon 2 [Figure 2]; two frameshift indels, exon 8), sloths (1-bp deletion, exon 3; 4-bp insertion, exon 7), and armadillos (1-bp deletion, exon 6), respectively. We were unable to obtain *MTNR1A* in most xenarthrans, possibly due to a whole gene deletion, although the assemblies were not complete enough to verify this via a synteny analysis. Nonetheless, the silky anteater (*Cyclopes didactylus*) and sloths retain one or more *MTNR1A* pseudogenes (Figure 2). Sloths appear to have multiple paralogs, the identities of which are difficult to tease out, with separate exons being found on separate contigs, but the pattern suggests all are probably inactivated. For *MTNR1B*, a 2-bp deletion in exon 2 is shared by sloths and anteaters (Pilosa; Figure 2). Among armadillos, a single 1-bp deletion (exon 1) is shared by dasypodids, and five inactivating mutations are shared among chlamyphorid armadillos (two in exon 1; three in exon 2).

From dN/dS ratio analyses, we found evidence of shifts in selection pressures consistent with ancient pseudogenization of these genes. Estimates in which key xenarthran branches had statistically elevated ω values were found for *AANAT* (background: 0.36; stem Folivora: 1.57*; crown Folivora: 0.71; stem Vermilingua: 2.81*; crown Vermilingua: 1.71*; stem Cingulata: 7.33*; crown Cingulata: 1.34*), *ASMT* (background: 0.25; crown Folivora: 0.56; crown Vermilingua: 1.07*; crown Cingulata: 0.91*; stem Cingulata: 1.06*), *MTNR1A* (background: 0.17; crown *Choloepus*: 1.06; stem *Choloepus*: 0.76*; *Cyclopes didactylus*: 1.11*) and *MTNR1B* (background: 0.32; crown Pilosa: 1.18*; crown Chlamyphoridae: 0.79*; stem Dasypodidae: 2.36*; * indicates model comparison remains statistically significant after Holm-Bonferroni correction, here and below). We also ran a model for *ASMT* in which crown Xenarthra was given a single ω, under the assumption that it was inactivated in the stem lineage (see above). This ω was estimated as 0.79, and when compared to a model in which the crown Xenarthra ω was fixed at 1, the former was a better fit for the data than the latter. As such, this result seems to be inconsistent with a stem xenarthran inactivation of *ASMT*.

Pholidota

For pangolins, all three species share two premature stop codons in *AANAT* (exons 1 and 2), another in exon 5 of *ASMT*, and four inactivating mutations in *MTNR1B* (Figure 2). Exon 1 of *MTNR1A* is missing for both *Manis* spp., preventing comparisons with *Phataginus tricuspis*, but the former two share a 1-bp deletion in exon 2 (Figure 2). Moreover, dN/dS analyses provide further evidence of pseudogenization with elevated ω values for *AANAT* (background: 0.36; crown Pholidota: 0.62), *ASMT* (background: 0.25; crown Pholidota: 0.72*), *MTNR1A* (background: 0.17; crown *Manis*: 1.11*; stem *Manis*: 3.35*; *Phataginus*: 1.43*; stem Pholidota: 0.35), and *MTNR1B* (background: 0.32; crown Pholidota: 1.51*; stem Pholidota: 1.14*).

Cetacea

Inactivation of the melatonin genes in cetaceans has already been reported in two recent studies (Huelsmann *et al.*, 2019; Lopes-Marques *et al.*, 2019), although we have expanded on these results by increasing the taxon sampling from 12 to 25 species and eight to 13 families (Figure 1), as well as analyzing the selection patterns of this clade. As in previous studies, we found shared inactivating mutations, including a premature stop codon in exon 1 of *AANAT* (all 25 species), a 1-bp insertion in exon 1 of *ASMT* (21 species, mysticetes + odontocetes), and a 1-bp deletion in exon 2 of *MTNR1B* (17 species, mysticetes + odontocetes) (Figure 2). For *MTNR1A*, almost all toothed whales (Odontoceti) completely lack the gene in their assemblies, with the exceptions of pseudogenes in *Platanista minor* and *Physeter catodon*, whereas all examined baleen whales (Mysticeti), *Platanista* and *Physeter* lack exon 2. Using contiguous assemblies from one mysticete and five odontocetes, we confirmed via a synteny analysis that exon 2 is absent (Extended data, Supplementary Figure S8 [Emerling *et al.*, 2021]), suggesting a whole exon deletion occurred on the stem Cetacea branch. Furthermore, an unusual 21-bp deletion in exon 1 shared by odontocetes and mysticetes may represent an in-frame disabling mutation (Figure 2). dN/dS analyses are consistent with relaxed selection in this clade, with statistically elevated ω values in *AANAT* (background: 0.36; crown Cetacea: 0.99*; stem Cetacea: 3.2), *ASMT* (background: 0.25; crown Cetacea: 0.93*; stem Cetacea: 0.7), *MTNR1A* (background: 0.17; stem Cetacea: 1.33) and *MTNR1B* (background: 0.32; crown Cetacea: 1.02*; stem Cetacea: 0.64).

Sirenia

For sirenians, shared inactivating mutations are less common, existing only for *MTNR1B* (exon 2: 14-bp deletion, 17-bp deletion [Figure 2], ancestral stop mutation TAC or TGC). Outside of this gene, inactivating mutations are only found in the manatee (*AANAT*: two; *ASMT*: one; *MTNR1A*: two) and dugong (*AANAT*: two; *MTNR1A*: two) (Figure 2), although the sole inactivating mutation in the manatee *ASMT* appears to be polymorphic based on short read data (Extended data, Supplementary Figure S12 [Emerling *et al.*, 2021]). Some of the

manatee mutations were previously reported (Huelsmann *et al.*, 2019; Lopes-Marques *et al.*, 2019). Here, dN/dS analyses only provide evidence of relaxed selection in *MTNR1A* (background: 0.17; stem Sirenia: 0.48; *Dugong*: 0.53*; *Trichechus*: 1.95*) and *MTNR1B* (background: 0.32; crown Sirenia: 1.95*).

Dermoptera

Both melatonin synthesis genes have shared inactivating mutations across all three colugos: exon 1 of *AANAT* and exon 3 of *ASMT* each have shared 1-bp deletions, (Figure 2). We were unable to obtain *MTNR1A* for *Cynocephalus volans*, but both *Galeopterus* species have a splice donor mutation (AG to AT) in the intron (Figure 2). Similarly, we were unable to assemble exon 2 of *MTNR1B* for *C. volans*, and exon 1 appears intact, but *Galeopterus* spp. share a 1-bp insertion (exon 1; Figure 2) and a premature stop codon (exon 2). Furthermore, dN/dS estimates provide evidence of relaxed selection in *AANAT* (background: 0.36; crown Dermoptera: 0.77), *ASMT* (background: 0.25; crown Dermoptera: 0.6) and *MTNR1A* (background: 0.17; crown *Galeopterus*: 0.81*; stem *Galeopterus*: 0.64*).

Other placental mammals

Beyond these taxa, *AANAT* is present as two to three paralogs in multiple non-cetacean cetartiodactyls, with one paralog sometimes being a pseudogene, a finding corroborated by a recent study (Yin *et al.*, 2021). However, at least one gene is always intact in all non-cetacean cetartiodactyls that we examined. *ASMT* is a pseudogene in three subterranean rodents (*Fukomys damarensis* [Bathyergidae]; *Nannospalax galili*, *Spalax carmeli* [Spalacidae]), *MTNR1A* is inactivated in two hyraxes (*Procavia capensis*, *Heterohyrax brucei*, Hyracoidea), a monk seal (*Neomonachus schauinslandi*), pig (*Sus scrofa*), talpid mole (*Condylura cristata*), and the naked-mole rat (*Heterocephalus glaber*), and *MTNR1B* is a pseudogene or likely deleted (i.e., negative BLAST results) in a host of other species, including seven afrotherians, eight carnivorans, one bat, all six examined eulipotyphlans, one lagomorph, five primates, and five rodents we examined (Figure 1). The few examples of shared inactivating mutations we found in these other species include a 4-bp insertion in exon 6 of *ASMT* in two spalacids (Extended data, Supplementary Figure S9 [Emerling *et al.*, 2021]), a 1-bp deletion in exon 2 of *MTNR1A* (Extended data, Supplementary Figure S10 [Emerling *et al.*, 2021]) and three inactivating mutations in *MTNR1B* shared by two hyraxes, and a 1-bp insertion in exon 2 of *MTNR1B* for two platyrrhine monkeys (*Callithrix jacchus*, *Cebus capucinus*). Here also, dN/dS ratio analyses suggest that many of these pseudogenes are indeed under relaxed selection based on statistically elevated ω values.

Discussion

Melatonin genes inactivated in many mammals, but functional in crocodylians

Here we reported evidence that xenarthrans, pangolins, cetaceans and some sirenians have lost the capability to synthesize and bind melatonin via the traditional pathway found in vertebrates, coinciding with the ostensible absence of a pineal gland in these taxa. This builds upon recent studies on cetaceans and the West Indian manatee (Huelsmann *et al.*, 2019; Lopes-Marques *et al.*, 2019), demonstrating the surprising extent of the degradation of these genes. We hypothesized that despite the pineal gland's apparent absence in xenarthrans, pangolins, sirenians, cetaceans and crocodylians, the genes underlying the production and signaling of melatonin would remain intact, given the widespread effects of melatonin in vertebrates and evidence of circulating melatonin in these taxa (Firth *et al.*, 2010; Funasaka *et al.*, 2011; Harlow *et al.*, 1981; Panin *et al.*, 2012; Roth *et al.*, 1980). Indeed, this appears to be the case for crocodylians, in which all 13 species we investigated, representing all major lineages, possess intact orthologs of *AANAT*, *ASMT*, *MTNR1A* and *MTNR1B*. Given the presence of serum melatonin in at least some crocodylians (Firth *et al.*, 2010; Roth *et al.*, 1980), this suggests that either the pineal gland is intact but difficult to isolate and/or extra-pineal sources of melatonin exist. In support of the latter hypothesis, we found evidence of *AANAT* and *ASMT* expression in the eyes of the American alligator (*Alligator mississippiensis*). Furthermore, a previous study found that the gene encoding a pineal opsin pigment is a pseudogene in crocodylians (Emerling, 2017b), potentially revealing a shift in the source of melatonin from the pineal gland to the eye, following the regression of the former.

By contrast, most major melatonin pathway genes are pseudogenized in pineal-less mammals. Researchers have varied in their reports on the presence versus absence of pineal glands in cetaceans (Behrmann, 1990; Holzmann, 1991; Lyamin *et al.*, 2008; McFarland *et al.*, 1969; Oelschläger *et al.*, 2008; Panin *et al.*, 2012) and work on sirenians appears to be inconsistent, with some studies suggestive of a minute pineal gland and others of its complete absence (Chapman, 1875; Murie, 1872; Ralph *et al.*, 1985). A recent anatomical study failed to find a pineal gland in the white-bellied pangolin (*Phataginus tricuspis*), although the absence was attributed to an error in preparing the brain for study (Imam *et al.*, 2019). Despite these inconsistencies, a regressed, if not completely absent, pineal gland largely predicts the disappearance of the canonical melatonin pathways, at least in mammals. Furthermore, we found evidence of melatonin synthesis inactivation in colugos, suggesting that these species may also possess a regressed pineal gland, or may lack it entirely. Although one study found a diencephalon in colugos comparable in relative size to gliding rodents and bats (Pirlot & Kamiya, 1982), a review of pinealocytes in mammals noted the lack of research specifically on the pineal in Dermoptera (Bhatnagar, 1992).

While extra-pineal melatonin may reconcile the patterns of circulating melatonin and the putative absence of a pineal gland in crocodylians, it cannot explain this same phenomenon in the nine-banded armadillo (Harlow *et al.*, 1982; Harlow *et al.*, 1981) and bottlenose dolphins (Funasaka *et al.*, 2011; Panin *et al.*, 2012). Possible sources for the serum melatonin detected in these taxa may derive from an unknown alternative pathway (Tan *et al.*, 2016), dietary sources (Hattori *et al.*, 1995; Peuhkuri *et al.*, 2012; Reiter *et al.*, 2005), or the organism's microbiome (Hardeland & Poeggeler, 2003). However, the functional significance of serum melatonin may be obviated by the

absence of melatonin receptors in these species, presumably preventing any contribution to circadian signaling in the body's tissues. Despite this, given that recent discoveries have shown that presenting as a pseudogene does not always indicate that all biological function is lost (Cheetham *et al.,* 2020), there remains a possibility that these apparently dysfunctional genes are able to contribute to melatonin metabolism in some unknown fashion.

Ancient loss of melatonin genes in placental mammals

Our results strongly suggest that melatonin synthesis and signaling have been abolished within multiple placental mammal lineages for extensive periods of evolutionary time. The most ancient of these may be in Xenarthra, given our evidence of possible *ASMT* and *MTNR1A* inactivation on the stem branch. Crown Xenarthra arose roughly 68 million years ago (mya) (Gibb *et al*., 2016), potentially meaning that the loss of melatonin synthesis and possibly some signaling took place near the K/Pg boundary, when non-avian dinosaurs went extinct and placental mammals began to radiate. However, given some ambiguity in the evidence for stem inactivation of *ASMT* (see results), and the absence of sequences of *MTNR1A* for most species of xenarthrans, convergent loss remains a strong possibility. Despite this, additional shared mutations in *AANAT* and *ASMT* suggest that the components for melatonin synthesis were disrupted prior to the origin of armadillos (45 mya), sloths (31 mya), and anteaters (38 mya), respectively. In addition, *MTNR1B* was likely pseudogenized prior to the sloth / anteater split (59 mya), and the origin of chlamyphorid armadillos (37 mya). In pangolins, shared mutations in *AANAT*, *ASMT*, and *MTNR1B* suggest complete loss of melatonin synthesis and at least some melatonin signaling prior to the origin of this clade 25 mya (Meredith *et al*., 2011). For colugos, *AANAT* and *ASMT* were likely inactivated prior to the origin of crown Dermoptera, indicating the absence of melatonin synthesis for at least 15 million years (Mason *et al*., 2016). There are contrasting patterns among the aquatic taxa, with cetaceans having convincingly lost all four genes prior to their origin 37 mya (McGowen *et al*., 2020), whereas for sirenians, we only have positive evidence of *MTNR1B* inactivation in the stem lineage at least 42 mya (Springer *et al*., 2015). Subsequent parallel pseudogenization events occurred in other melatonin-related genes for the manatee and dugong, but not Steller's sea cow.

Causes and consequences of losing the melatonin pathway

The potential significance of the loss of both melatonin synthesis and signaling in multiple clades of mammals should not be understated. To reiterate, melatonin is a ubiquitous biogenic compound found in Eubacteria, unicellular eukaryotes, plants, fungi and animals (Hardeland & Poeggeler, 2003). While it is unclear if melatonin synthesis has a single origin or evolved independently in several lineages (Tan *et al*., 2014; Zhao *et al*., 2019), its widespread taxonomic occurrence suggests that it has ancient origins and that natural selection has favored the maintenance of synthesis pathways for perhaps billions of years. Furthermore, after a hypothesized co-option of this potent antioxidant to signal environmental darkness (Zhao *et al*., 2019), in order to help modulate circadian and circannual physiological processes, melatonin synthesis and signaling would seemingly become indispensable for most vertebrates. As such, it is a challenge to clarify the causes and consequences of losing melatonin pathway genes (Valente *et al.,* 2021).

Convergent evolution often results from similar selection pressures, which may explain why both cetaceans and some sirenians have lost these genes. Perhaps the unique demands of a fully aquatic lifestyle, such as the need to frequently surface for respiration, needed to be uncoupled from a rhythmic signal of darkness. By contrast, the semi-aquatic sea otter (*Enhydra lutris*) and pinnipeds have retained their melatonin synthesis genes, although *MTNR1B* is pseudogenized in *E. lutris* and two seals (Phocidae), with one of the phocids also showing evidence of *MTNR1A* inactivation. Perhaps this underlies their intermediate aquatic phenotype, although eight of the 13 carnivorans also present an *MTNR1B* pseudogene. Another example of strong convergent evolution can be seen in xenarthran anteaters and pangolins, both of which have radically modified their feeding apparatus to ingest ants and termites (myrmecophagy). However, other myrmecophagous taxa we examined, including the aardvark (*Orycteropus afer*), aardwolf (*Proteles cristatus*), and bat-eared fox (*Otocyon megalotis*) at most only have *MTNR1B* inactivated.

Regardless of their specific phenotypes, all of these taxa experience fluctuations in light and darkness, so it is unclear as to why loss of such a hormone would be beneficial. By contrast, it seems more logical for melatonin synthesis to be lost while adapting to an environment of nearly complete darkness, in which rhythmic secretions entrained on light patterns may no longer be possible. Multiple lineages of subterranean mammals fit this description, and indeed, we found evidence of *ASMT* pseudogenes in the subterranean rodents *Nannospalax galili*, *Spalax carmeli* (Spalacidae) and *Fukomys damarensis* (Bathyergidae), and inactivation of both receptors in the naked mole-rat (*Heterocephalus glaber*; Bathyergidae). The latter result had been previously reported in a single individual (Kim *et al*., 2011), but we confirmed that both genes share the same disabling mutations in a second individual *H. glaber* and are likely under relaxed selection. The fossorial star-nosed mole, *Condylura cristata* (Talpidae), also appears to have dispensed of both melatonin receptors, with *MTNR1A* a pseudogene and *MTNR1B* being completely absent from the assembly. In addition, dN/dS estimates suggest selection on *ASMT* is relaxed in *H. glaber*, a species which also has an atrophied pineal gland (Quay, 1981), and *MTNR1B* is inactivated in *N. galili*, *F. damarensis* and a golden mole (*Chrysochloris asiatica*; Chrysochloridae). This may be relevant to the ancient loss of melatonin synthesis and signaling in xenarthrans, given that comparative anatomy and an analysis of genes critical for vision in bright light appear to point to an early subterranean history for this clade (Emerling & Springer, 2015). Perhaps an extended history underground limited the utility of melatonin synthesis and signaling, and upon emerging from this committed existence in the darkness, their descendants inherited this unusual phenotype. However, a pineal gland that can synthesize melatonin has been reported in

at least one spalacid (Balemans *et al*., 1980), pineal glands are reported to be present in talpids and chrysochlorids (Legait *et al*., 1976; Pevet, 1974; Pevet & Kuyper, 1978) and the melatonin synthesis genes remain intact for *Chrysochloris asiatica* and *Condylura cristata*.

One potentially unifying concept for the pattern of pineal gland / melatonin synthesis loss may be related to thermoregulation. Ralph (1975) hypothesized that the size of the pineal gland in vertebrates may be correlated to thermoregulation, tentatively linking pineal gland size to latitude, activity pattern and relative homeothermy. He observed that some of the largest pineal glands belong to species that inhabit higher latitudes, while pointing out that some vertebrates with the smallest or absent pineal glands tend to be restricted to the tropics. Though the comparisons were limited, a better-substantiated pattern was noted with the related parietal eye of squamates. The parietal eye is an organ that is developmentally related to and anatomically linked with the pineal gland, which appears to largely provide information for thermoregulation in ectothermic squamates, possibly through melatonin regulation (Firth & Kennaway, 1980; Hutchison & Kosh, 1974; Phillips & Harlow, 1981; Stebbins & Eakin, 1958). When comparing the presence or absence of the parietal eye, researchers noted a trend of parietal eye loss in squamates that live near the equator (Gundy *et al*., 1975). Given that the parietal eye provides information about temperature and light, which largely correlates with the amount of sunlight, and the pineal gland secretes melatonin in darkness, the latitudinal hypothesis may have some validity. At lower latitudes, there is less seasonality; therefore, being able to detect changes in circadian and circannual dark and light cues is plausibly of less adaptive benefit in these regions. Notably, xenarthrans, pangolins, sirenians and colugos live almost exclusively in the tropics; furthermore, aquatic and subterranean habitats provide a buffering effect from temperature fluctuation. These characteristics encompass all taxa we record as lacking melatonin synthesis capabilities, making it a potentially attractive hypothesis.

Significantly, the patterns of melatonin pathway degradation have strong overlap with placental mammals that have lost the capacity for non-shivering thermogenesis (NST). Specifically, xenarthrans, pangolins, cetaceans and sirenians all have inactivated *UCP1*, a gene that facilitates NST (Gaudry *et al*., 2017). Furthermore, hyraxes (Hyracoidea) and pigs (Suidae) have a *UCP1* pseudogene, hyraxes (*Procavia capensis, Heterohyrax brucei*) have both melatonin receptor genes inactivated, and the wild boar (*Sus scrofa*) has an *MTNR1A* pseudogene. Notably, melatonin induces the production of brown adipose tissue (Heldmaier *et al*., 1981; Heldmaier & Hoffmann, 1974), a major location of NST. Together, these data suggest that the loss of melatonin synthesis may be coupled with the loss of this thermoregulatory tool in these clades, further underscoring a potential link between melatonin pathway loss and changes in thermoregulatory requirements.

Conclusion

In this study we provided evidence that, in contrast to crocodylians, numerous placental mammals reported to lack a pineal gland also lack the genes necessary for the canonical vertebrate melatonin synthesis and signaling pathways. However, this result seems to raise more questions than answers. Given the pleiotropic nature of melatonin synthesis and signaling genes, which selection pressure(s) could have led to the loss of this seemingly crucial signaling molecule? What are the physiological consequences of this loss? Are there possibly compensatory alternative mechanisms for producing and sensing melatonin? For those species that present serum melatonin, how are they doing so? Does this melatonin function merely as an antioxidant, or does it aid in circadian and circannual signaling via different pathways? Further studies on comparative anatomy, physiology and gene expression in pineal-less taxa and others should shed further light on these challenging questions.

Data availability

Underlying data

Zenodo: Genomic evidence for the parallel regression of melatonin synthesis and signaling pathways in placental mammals. https://doi.org/10.5281/zenodo.4894211 (Emerling *et al*., 2021)

This project contains the following underlying data:

- Supplementary_Dataset_S1.txt: Genomic alignments in fasta format used to determine the pseudogene/functional status of all four melatonin genes in different taxonomic groups.
- Supplementary_Dataset_S2.txt: Alignment of *AANAT* in phylip format used in maximum likelihood phylogenetic reconstruction with RAxML.
- Supplementary_Dataset_S3.txt: Alignment of *ASMT* in phylip format used in maximum likelihood phylogenetic reconstruction with RAxML.
- Supplementary_Dataset_S4.txt: Alignment of *MTNR1A* and *MTNR1B* in phylip format used in maximum likelihood phylogenetic reconstruction with RAxML.
- Supplementary_Dataset_S5.txt: Codon alignments of *AANAT* used in selection pressure analyses with PAML.
- Supplementary_Dataset_S6.txt: Codon alignments of *ASMT* used in selection pressure analyses with PAML.
- Supplementary_Dataset_S7.txt: Codon alignments of *MTNR1A* used in selection pressure analyses with PAML.
- Supplementary_Dataset_S8.txt: Codon alignments of *MTNR1B* used in selection pressure analyses with PAML.
- Supplementary_Dataset_S9.txt: Tree topologies in newick format used in selection pressure analyses with PAML.
- Supplementary_Table_S1.xlsx: List of species examined in this study and the sources of the genes.
- Supplementary_Table_S2.xlsx: Accession numbers and functionality of *AANAT* in species examined.

- Supplementary_Table_S3.xlsx: Accession numbers and functionality of *ASMT* in species examined.
- Supplementary_Table_S4.xlsx: Accession numbers and functionality of *MTNR1A* in species examined.
- Supplementary_Table_S5.xlsx: Accession numbers and functionality of *MTNR1B* in species examined.
- Supplementary_Table_S6.xlsx: Codon frequency model selection.
- Supplementary_Table_S7.xlsx: Results of *AANAT* PAML dN/dS analyses for mammals.
- Supplementary_Table_S8.xlsx: Results of *ASMT* PAML dN/dS analyses for mammals.
- Supplementary_Table_S9.xlsx: Results of *MTNR1A* PAML dN/dS analyses for mammals.
- Supplementary_Table_S10.xlsx: Results of *MTNR1B* PAML dN/dS analyses for mammals.
- Supplementary_Table_S11.xlsx: Results of PAML analyses for sauropsids.
- Supplementary_Table_S12.xlsx: Results of BLASTing and mapping short reads from *Alligator mississippiensis* RNA sequencing experiments.
- Supplementary_Table_S13.xlsx: Supporting data for validating putative inactivating mutations.

Extended data

Zenodo: Genomic evidence for the parallel regression of the melatonin synthesis and signaling pathways in placental mammals. http://doi.org/10.5281/zenodo.4894211 (Emerling *et al.*, 2021)

This project contains the following extended data:

- Supplementary_Figure_S1.pdf: *AANAT* PAML 'master model' showing branch categories, corresponding to "Model 1: 24 ratio" in Supplementary Table S7.
- Supplementary_Figure_S2.pdf: *ASMT* PAML 'master model' showing branch categories, corresponding to "Model 2: 24 ratio" in Supplementary Table S8.
- Supplementary_Figure_S3.pdf: *MTNR1A* PAML 'master model' showing branch categories, corresponding to "Model 1: 27 ratio" in Supplementary Table S9.
- Supplementary_Figure_S4.pdf: *MTNR1B* PAML 'master model' showing branch categories, corresponding to "Model 1: 46 ratio" in Supplementary Table S10.
- Supplementary_Figure_S5.pdf: RAxML *AANAT* gene tree. Numbers at nodes correspond to bootstrap support values.
- Supplementary_Figure_S6.pdf: RAxML *ASMT* gene tree. Numbers at nodes correspond to bootstrap support values.
- Supplementary_Figure_S7.pdf: RAxML *MTNR1A+MTNR1B* tree. Numbers at nodes correspond to bootstrap support values.
- Supplementary_Figure_S8.pdf: supporting data showing the inactivation of *MTNR1A* exon 2 in cetaceans.
- Supplementary_Figure_S9.pdf: supporting data showing the inactivation of *ASMT* in spalacids and *Fukomys damarensis*.
- Supplementary_Figure_S10.pdf: supporting data showing the inactivation of *MTNR1A* in hyracoids and *Cyclopes didactylus*.
- Supplementary_Figure_S11.pdf: supporting data showing the inactivation of *MTNR1A* in sirenians.
- Supplementary_Figure_S12.pdf: supporting data showing the inactivation of *AANAT* in sirenians and a polymorphic premature stop codon in exon 5 of *ASMT* in *Trichechus manatus*.
- Supplementary_Figure_S13.pdf: supporting data showing the inactivation of *MTNR1A* in *Condylura cristata*.
- Supplementary_Figure_S14.pdf: supporting data showing the inactivation of *MTNR1A* in *Phataginus tricuspis*.
- Supplementary_Figure_S15.pdf: PAML *AANAT* results, Model 1: 24 ratio.
- Supplementary_Figure_S16.pdf: PAML *ASMT* results, Model 2: 24 ratio.
- Supplementary_Figure_S17.pdf: PAML *MTNR1A* results, Model 1: 27 ratio.
- Supplementary_Figure_S18.pdf: PAML *MTNR1B* results, Model 1: 46 ratio.

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Acknowledgements

We thank Drs. Mareike C. Janiak and Ryan K. Schott for their helpful contributions in reviewing an earlier version of this manuscript. We thank G. Amato for usage of unpublished genomic sequences for two crocodylian species and the St. Augustine Alligator Farm Zoological Park (St. Augustine, Florida, USA) for tissue samples of crocodylians. This is contribution ISEM 2021-149 of the Institut des Sciences de l'Evolution de Montpellier.

References

Acuña-Castroviejo D, Escames G, Venegas C, *et al.*: **Extrapineal melatonin: Sources, regulation, and potential functions.** *Cell Mol Life Sci.* 2014; **71**(16): 2997–3025.
PubMed Abstract | **Publisher Full Text**

Albalat R, Cañestro C: **Evolution by gene loss.** *Nat Rev Genet.* 2016; **17**(7): 379–391.
PubMed Abstract | **Publisher Full Text**

Allio R, Tilak MK, Scornavacca C, *et al.*: **High-quality carnivoran genomes from roadkill samples enable comparative species delineation in aardwolf and bat-eared fox.** *eLife.* 2021; **10**: e63167.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Baburski AZ, Sokanovic SJ, Janjic MM, *et al.*: **Melatonin replacement restores the circadian behavior in adult rat Leydig cells after pinealectomy.** *Mol Cell Endocrinol.* 2015; **413**: 26–35.
PubMed Abstract | **Publisher Full Text**

Balemans MG, Pévet P, Legerstee WC, *et al.*: **Preliminary investigations on melatonin and 5-methoxytryptophol synthesis in the pineal, retina, and Harderian gland of the mole rat and in the pineal of the mouse "eyeless."** *J Neural Transm.* 1980; **49**(4): 247–255.
PubMed Abstract | **Publisher Full Text**

Behrmann G: **The pineal organ (Epiphysis cerebri) of the harbour porpoise *Phocoena phocoena*. (Linné, 1758).** *Aquat Mamm.* 1990; **16**(3): 96–100.
Reference Source

Berger MJ, Wenger AM, Guturu H, *et al.*: **Independent erosion of conserved transcription factor binding sites points to shared hindlimb, vision and external testes loss in different mammals.** *Nucleic Acids Res.* 2018; **46**(18): 9299–9308.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Bhatnagar KP: **The ultrastructure of mammalian pinealocytes: A systematic investigation.** *Microsc Res Tech.* 1992; **21**(2): 85–115.
PubMed Abstract | **Publisher Full Text**

Bittman EL: **The sites and consequences of melatonin binding in mammals.** *Am Zoo.* 1993; **33**(2): 200–211.
Publisher Full Text

Canpolat S, Sandal S, Yilmaz B, *et al.*: **Effects of pinealectomy and exogenous melatonin on serum leptin levels in male rat.** *Eur J Pharmacol.* 2001; **428**(1): 145–148.
PubMed Abstract | **Publisher Full Text**

Chapman HC: **Observations on the structure of the manatee.** *Proc Acad Nat Sci Philadelphia.* 1875; **27**(3): 452–462.
Reference Source

Cheetham SW, Faulkner GJ, Dinger ME: **Overcoming challenges and dogmas to understand the functions of pseudogenes.** *Nat Rev Genet.* 2020; **21**(3): 191–201.
PubMed Abstract | **Publisher Full Text**

Cipolla-Neto J, Amaral FG, Afeche SC, *et al.*: **Melatonin, energy metabolism, and obesity: A review.** *J Pineal Res.* 2014; **56**(4): 371–381.
PubMed Abstract | **Publisher Full Text**

Cipolla-Neto J, Do Amaral FG: **Melatonin as a hormone: New physiological and clinical insights.** *Endocr Rev.* 2018; **39**(6): 990–1028.
PubMed Abstract | **Publisher Full Text**

Coelho LA, Peres R, Amaral FG, *et al.*: **Daily differential expression of melatonin-related genes and clock genes in rat cumulus-oocyte complex: Changes after pinealectomy.** *J Pineal Res.* 2015; **58**(4): 490–499.
PubMed Abstract | **Publisher Full Text**

Cogburn LA, Wilson-Placentra S, Letcher LR: **Influence of pinealectomy on plasma and extrapineal melatonin rhythms in young chickens (*Gallus domesticus*).** *Gen Comp Endocrinol.* 1987; **68**(3): 343–356.
PubMed Abstract | **Publisher Full Text**

David-Gray ZK, Bellingham J, Munoz M, *et al.*: **Adaptive loss of ultraviolet-sensitive/violet-sensitive (UVS/VS) cone opsin in the blind mole rat (*Spalax ehrenbergi*).** *Eur J Neurosci.* 2002; **16**(7): 1186–1194.
PubMed Abstract | **Publisher Full Text**

De Farias TDS, De Oliveira AC, Andreotti S, *et al.*: **Pinealectomy interferes with the circadian clock genes expression in white adipose tissue.** *J Pineal Res.* 2015; **58**(3): 251–261.
PubMed Abstract | **Publisher Full Text**

Del Gobbo V, Libri V, Villani N, *et al.*: **Pinealectomy inhibits interleukin-2 production and natural killer activity in mice.** *Int J Immunopharmacol.* 1989; **11**(5): 567–573.
PubMed Abstract | **Publisher Full Text**

Deméré TA, McGowen MR, Berta A, *et al.*: **Morphological and molecular evidence for a stepwise evolutionary transition from teeth to baleen in mysticete whales.** *Syst Biol.* 2008; **57**(1): 15–37.
PubMed Abstract | **Publisher Full Text**

Edgar RC: **MUSCLE: Multiple sequence alignment with high accuracy and high throughput.** *Nucleic Acids Res.* 2004; **32**(5): 1792–1797.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Emerling CA: **Regressed but not gone: Patterns of vision gene loss and retention in subterranean mammals.** *Integr Comp Biol.* 2018; **58**(3): 441–451.
PubMed Abstract | **Publisher Full Text**

Emerling CA: **Genomic regression of claw keratin, taste receptor and light-associated genes provides insights into biology and evolutionary origins of snakes.** *Mol Phylogenet Evol.* 2017a; **115**: 40–49.
PubMed Abstract | **Publisher Full Text**

Emerling CA: **Archelosaurian color vision, parietal eye loss, and the crocodylian nocturnal bottleneck.** *Mol Biol Evol.* 2017b; **34**(3): 666–676.
PubMed Abstract | **Publisher Full Text**

Emerling CA, Huynh HT, Nguyen MA, *et al.*: **Spectral shifts of mammalian ultraviolet-sensitive pigments (Short wavelength-sensitive opsin 1) are associated with eye length and photic niche evolution.** *Proc Biol Sci.* 2015; **282**(1819): 20151817.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Emerling CA, Springer MS: **Genomic evidence for rod monochromacy in sloths and armadillos suggests early subterranean history for Xenarthra.** *Proc Biol Sci.* 2015; **282**(1800): 20142192.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Emerling CA, Springer MS: **Eyes underground: Regression of visual protein networks in subterranean mammals.** *Mol Phylogenet Evol.* 2014; **78**: 260–270.
PubMed Abstract | **Publisher Full Text**

Emerling CA, Springer MS, Gatesy J, *et al.*: **Genomic evidence for the parallel regression of the melatonin pathway in placental mammals (Version 1).** [Data set]. Open Research Europe. *Zenodo*. 2021.
http://www.doi.org/10.5281/zenodo.4894211

Fang X, Nevo E, Han L, *et al.*: **Genome-wide adaptive complexes to underground stresses in blind mole rats *Spalax*.** *Nat Commun.* 2014; **5**: 3966.
PubMed Abstract | **Publisher Full Text**

Firth BT, Christian KA, Belan I, *et al.*: **Melatonin rhythms in the Australian freshwater crocodile (*Crocodylus johnstoni*): a reptile lacking a pineal complex?** *J Comp Physiol B.* 2010; **180**(1): 67–72.
PubMed Abstract | **Publisher Full Text**

Firth BT, Kennaway DJ: **Plasma melatonin levels in the scincid lizard *Trachydosaurus rugosus*: The effects of parietal eye and lateral eye impairment.** *J Exp Biol.* 1980; **85**(1): 311–321.
Publisher Full Text

Fjelldal PG, Grotmol S, Kryvi H, *et al.*: **Pinealectomy induces malformation of the spine and reduces the mechanical strength of the vertebrae in Atlantic salmon, *Salmo salar*.** *J Pineal Res.* 2004; **36**(2): 132–139.
PubMed Abstract | **Publisher Full Text**

Foà A, Menaker M: **Contribution of pineal and retinae to the circadian rhythms of circulating melatonin in pigeons.** *J Comp Physiol A.* 1988; **164**: 25–30.
Publisher Full Text

Funasaka N, Yoshioka M, Suzuki M, *et al.*: **Seasonal difference of diurnal variations in serum melatonin, cortisol, testosterone, and rectal temperature in Indo-Pacific bottlenose dolphins (*Tursiops aduncus*).** *Aquat Mamm.* 2011; **37**(4): 433–442.
Publisher Full Text

Gaetano J: **Holm-Bonferroni sequential correction: An Excel Calculator (1.3).** [Microsoft Excel workbook]. 2018.
Publisher Full Text

Gaudry MJ, Jastroch M, Treberg JR, *et al.*: **Inactivation of thermogenic UCP1 as a historical contingency in multiple placental mammal clades.** *Sci Adv.* 2017; **3**(7): e1602878.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Gern WA, Norris DO: **Plasma melatonin in the neotenic tiger salamander (*Ambystoma tigrinum*): Effects of photoperiod and pinealectomy.** *Gen Comp Endocrinol.* 1979; **38**(4): 393–398.
PubMed Abstract | **Publisher Full Text**

Gern WA, Owens DW, Ralph CL: **The synthesis of melatonin by the trout retina.** *J Exp Zool.* 1978a; **206**(2): 263–270.
Publisher Full Text

Gern WA, Owens DW, Ralph CL: **Persistence of the nychthemeral rhythm of melatonin secretion in pinealectomized or optic tract-sectioned trout (*Salmo gairdneri*).** *J Exp Zool.* 1978b; **205**(3): 371–376.
Publisher Full Text

Gibb GC, Condamine FL, Kuch M, *et al.*: **Shotgun mitogenomics provides a reference phylogenetic framework and timescale for living xenarthrans.** *Mol Biol Evol.* 2016; **33**(3): 621–642.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Gundy GC, Ralph CL, Wurst GZ: **Parietal eyes in lizards: Zoogeographical correlates.** *Science.* 1975; **190**(4215): 671–673.
PubMed Abstract | **Publisher Full Text**

Haldar C, Ghosh M: **Annual pineal and testicular cycle in the indian jungle bush quail, *Perdicula asiatica*, with reference to the effect of pinealectomy.** *Gen Comp Endocrinol.* 1990; **77**(1): 150–157.
PubMed Abstract | **Publisher Full Text**

Hardeland R, Poeggeler B: **Non-vertebrate melatonin.** *J Pineal Res.* 2003; **34**(4):

233–241.
PubMed Abstract | **Publisher Full Text**

Harlow HJ, Phillips JA, Ralph CL: **Circadian rhythms and the effects of exogenous melatonin in the nine-banded armadillo, *Dasypus novemcinctus*: A mammal lacking a distinct pineal gland.** *Physiol Behav.* 1982; **29**(2): 307–313.
PubMed Abstract | **Publisher Full Text**

Harlow HJ, Phillips JA, Ralph CL: **Day - night rhythm in plasma melatonin in a mammal lacking a distinct pineal gland, the nine-banded armadillo.** *Gen Comp Endocrinol.* 1981; **45**(2): 212–218.
PubMed Abstract | **Publisher Full Text**

Hattori A, Migitaka H, Iigo M, *et al.*: **Identification of melatonin in plants and its effects on plasma melatonin levels and binding to melatonin receptors in vertebrates.** *Biochem Mol Biol Int.* 1995; **35**(3): 627–34.
PubMed Abstract

He K, Shinohara A, Helgen KM, *et al.*: **Talpid mole phylogeny unites shrew moles and illuminates overlooked cryptic species diversity.** *Mol Biol Evol.* 2017; **34**(1): 78–87.
PubMed Abstract | **Publisher Full Text**

Hekkala E, Gatesy J, Narechania A, *et al.*: **Paleogenomics illuminates the evolutionary history of the extinct Holocene "horned" crocodile of Madagascar, *Voay robustus*.** *Commun Biol.* 2021; **4**(1): 505.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Heldmaier G, Hoffmann K: **Melatonin stimulates growth of brown adipose tissue.** *Nature.* 1974; **247**(5438): 224–225.
PubMed Abstract | **Publisher Full Text**

Heldmaier G, Steinlechner S, Rafael J, *et al.*: **Photoperiodic control and effects of melatonin on nonshivering thermogenesis and brown adipose tissue.** *Science.* 1981; **212**(4497): 917–919.
PubMed Abstract | **Publisher Full Text**

Holm S: **A simple sequential rejective method procedure.** *Scand J Stat.* 1979; **6**(2): 65–70.
Reference Source

Holzmann T: **Morphologie und mikroskopische Anatomie des Gehirns beim fetalen Narwal, *Monodon monoceros*.** University of Frankfurt am Main, 1991.

Huelsmann M, Hecker N, Springer MS, *et al.*: **Genes lost during the transition from land to water in cetaceans highlight genomic changes involved in aquatic adaptations.** *bioRxiv.* 2019.
Publisher Full Text

Hutchison VH, Kosh RJ: **Thermoregulatory function of the parietal eye in the lizard *Anolis carolinensis*.** *Oecologia.* 1974; **16**(2): 173–177.
PubMed Abstract | **Publisher Full Text**

Imam A, Bhagwandin A, Ajao MS, *et al.*: **The brain of the tree pangolin (*Manis tricuspis*). V. The diencephalon and hypothalamus.** *J Comp Neurol.* 2019; **527**(15): 2413–2439.
PubMed Abstract | **Publisher Full Text**

Kearse M, Moir R, Wilson A, *et al.*: **Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data.** *Bioinformatics.* 2012; **28**(12): 1647–1649.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Kennaway DJ, Frith RG, Phillipou G, *et al.*: **A specific radioimmunoassay for melatonin in biological tissue and fluids and its validation by gas chromatography-mass spectrometry.** *Endocrinology.* 1977; **101**(1): 119–127.
PubMed Abstract | **Publisher Full Text**

Khaldy H, León J, Escames G, *et al.*: **Circadian rhythms of dopamine and dihydroxyphenyl acetic acid in the mouse striatum: Effects of pinealectomy and of melatonin treatment.** *Neuroendocrinology.* 2002; **75**(3): 201–208.
PubMed Abstract | **Publisher Full Text**

Kim EB, Fang X, Fushan AA, *et al.*: **Genome sequencing reveals insights into physiology and longevity of the naked mole rat.** *Nature.* 2011; **479**(7372): 223–227.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Kumar S, Stecher G, Suleski M, *et al.*: **TimeTree: A Resource for Timelines, Timetrees, and Divergence Times.** *Mol Biol Evol.* 2017; **34**(7): 1812–1819.
PubMed Abstract | **Publisher Full Text**

Kvon EZ, Kamneva OK, Melo US, *et al.*: **Progressive loss of function in a limb enhancer during snake evolution.** *Cell.* 2016; **167**(3): 633–642.e11.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Legait H, Bauchot R, Stephan H, *et al.*: **Étude des corrélations liant le volume de l'épiphyse aux poids somatique et encéphalique chez les rongeurs, les insectivores, les chiropteres, les prosimiens et les simiens.** *Mammalia.* 1976; **40**(2): 327–337.
Publisher Full Text

Lima FB, Machado UF, Bartol I, *et al.*: **Pinealectomy causes glucose intolerance and decreases adipose cell responsiveness to insulin in rats.** *Am J Physiol.* 1998; **275**(6): E934–41.
PubMed Abstract | **Publisher Full Text**

Lopes-Marques M, Ruivo R, Alves LQ, *et al.*: **The singularity of Cetacea behavior parallels the complete inactivation of melatonin gene modules.** *Genes (Basel).* 2019; **10**(2): 121.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Lyamin OI, Manger PR, Ridgway SH, *et al.*: **Cetacean sleep: An unusual form of mammalian sleep.** *Neurosci Biobehav Rev.* 2008; **32**(8): 1451–1484.
PubMed Abstract | **Publisher Full Text**

Lynch HJ, Ozaki Y, Shakal D, *et al.*: **Melatonin excretion of man and rats: Effect of time of day, sleep, pinealectomy and food consumption.** *Int J Biometeorol.* 1975; **19**(4): 267–279.
PubMed Abstract | **Publisher Full Text**

Mason VC, Li G, Minx P, *et al.*: **Genomic analysis reveals hidden biodiversity within colugos, the sister group to primates.** *Sci Adv.* 2016; **2**(8): e1600633.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

McFarland WL, Morgane PJ, Jacobs MS: **Ventricular system of the brain of the dolphin, *Tursiops truncatus*, with comparative anatomical observations and relations to brain specializations.** *J Comp Neurol.* 1969; **135**(3): 275–368.
PubMed Abstract | **Publisher Full Text**

McGowen MR, Tsagkogeorga G, Álvarez-Carretero S, *et al.*: **Phylogenomic resolution of the cetacean tree of life using target sequence capture.** *Syst Biol.* 2020; **69**(3): 479–501.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Meredith RW, Gatesy J, Murphy WJ, *et al.*: **Molecular decay of the tooth gene enamelin (*ENAM*) mirrors the loss of enamel in the fossil record of placental mammals.** *PLoS Genet.* 2009; **5**(9): e1000634.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Meredith RW, Janečka JE, Gatesy J, *et al.*: **Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification.** *Science.* 2011; **334**(6055): 521–524.
PubMed Abstract | **Publisher Full Text**

Meredith RW, Zhang G, Gilbert MTP, *et al.*: **Evidence for a single loss of mineralized teeth in the common avian ancestor.** *Science.* 2014; **346**(6215): 1254390.
PubMed Abstract | **Publisher Full Text**

Miller MA, Pfeiffer W, Schwartz T: **Creating the CIPRES Science Gateway for inference of large phylogenetic trees.** *2010 Gatew Comput Environ Work GCE.* 2010.
Publisher Full Text

Murie J: **V. On the form and structure of the manatee (*Manatus americanus*).** *Trans Zool Soc London.* 1872; **8**(3): 127–220.
Publisher Full Text

Nishiwaki-Ohkawa T, Yoshimura T: **Molecular basis for regulating seasonal reproduction in vertebrates.** *J Endocrinol.* 2016; **229**(3): R117–R127.
PubMed Abstract | **Publisher Full Text**

Oelschläger HHA, Haas-Rioth M, Fung C, *et al.*: **Morphology and evolutionary biology of the dolphin (*Delphinus* sp.) brain--MR imaging and conventional histology.** *Brain Behav Evol.* 2008; **71**(1): 68–86.
PubMed Abstract | **Publisher Full Text**

Oksche A: **Survey of the development and comparative morphology of the pineal organ.** *Prog Brain Res.* In: *Structure and Function of the Epiphysis Cerebri.* Elsevier Publishing Company, 1965; **10**: 3–29.
PubMed Abstract | **Publisher Full Text**

Owens DW, Gern WA: **The pineal gland and melatonin in sea turtles.** In: Lofts, B., Holmes, W.N. (Eds.), *Current Trends in Comparative Endocrinology.* Hong Kong University Press, Hong Kong, 1985; 645–648.

Panin M, Gabai G, Ballarin C, *et al.*: **Evidence of melatonin secretion in cetaceans: Plasma concentration and extrapineal HIOMT-like presence in the bottlenose dolphin *Tursiops truncatus*.** *Gen Comp Endocrinol.* 2012; **177**(2): 238–245.
PubMed Abstract | **Publisher Full Text**

Partha R, Chauhan BK, Ferreira Z, *et al.*: **Subterranean mammals show convergent regression in ocular genes and enhancers, along with adaptation to tunneling.** *eLife.* 2017; **6**: e25884.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Pereira AG, Sterli J, Moreira FRR, *et al.*: **Multilocus phylogeny and statistical biogeography clarify the evolutionary history of major lineages of turtles.** *Mol Phylogenet Evol.* 2017; **113**: 59–66.
PubMed Abstract | **Publisher Full Text**

Peuhkuri K, Sihvola N, Korpela R: **Dietary factors and fluctuating levels of melatonin.** *Food Nutr Res.* 2012; **56**: 17252.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Pevet P: **The pineal gland of the mole (*Talpa europaea* L.) I. The fine structure of the pinealocytes.** *Cell Tissue Res.* 1974; **153**(3): 277–292.
PubMed Abstract | **Publisher Full Text**

Pevet P, Kuyper MA: **The ultrastructure of pinealocytes in the golden mole (*Amblysomus hottentotus*) with special reference to the granular vesicles.** *Cell Tissue Res.* 1978; **191**(1): 39–56.
PubMed Abstract | **Publisher Full Text**

Phillips JA, Harlow HJ: **Elevation of upper voluntary temperatures after shielding the parietal eye of horned lizards (*Phrynosoma douglassi*).** *Herpetologica.* 1981; **37**(4): 199–205.
Reference Source

Phillips JA, Harlow HJ, McArthur NH, *et al.*: **Epithalamus of the nine-banded armadillo, *Dasypus novemcinctus*.** *Comp Biochem Physiol A Comp Physiol.* 1986; **85**(3): 477–481.
PubMed Abstract | **Publisher Full Text**

Pirlot P, Kamiya T: **Relative size of brain and brain components in three gliding placentals (Dermoptera: Rodentia).** *Can J Zool.* 1982; **60**(4): 565–572.
Publisher Full Text

Prum RO, Berv JS, Dornburg A, *et al.*: **A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing.** *Nature.* 2015; **526**(7574): 569–573.
PubMed Abstract | **Publisher Full Text**

Quay WB: **Pineal atrophy and other neuroendocrine and circumventricular features of the naked mole-rat, *Heterocephalus glaber* (Rüppell), a fossorial, equatorial rodent.** *J Neural Transm.* 1981; **52**(1–2): 107–115.
PubMed Abstract | **Publisher Full Text**

Ralph CL: **The pineal gland and geographical distribution of animals.** *Int J Biometeorol.* 1975; **19**(4): 289–303.
PubMed Abstract | **Publisher Full Text**

Ralph CL, Young S, Gettinger R, *et al.*: **Does the manatee have a pineal body?** *Acta Zool.* 1985; **66**(1): 55–60.
Publisher Full Text

Reiter RJ, Manchester LC, Tan DX: **Melatonin in walnuts: Influence on levels of melatonin and total antioxidant capacity of blood.** *Nutrition.* 2005; **21**(9): 920–924.
PubMed Abstract | **Publisher Full Text**

Reppert SM, Sagar SM: **Characterization of the day-night variation of retinal melatonin content in the chick.** *Invest Ophthalmol Vis Sci.* 1983; **24**(3): 294–300.
PubMed Abstract

Roth JJ, Gern WA, Roth EC, *et al.*: **Nonpineal melatonin in the alligator (*Alligator mississippiensis*).** *Science.* 1980; **210**(4469): 548–550.
PubMed Abstract | **Publisher Full Text**

Sackton TB, Grayson P, Cloutier A, *et al.*: **Convergent regulatory evolution and loss of flight in paleognathous birds.** *Science.* 2019; **364**(6435): 74–78.
PubMed Abstract | **Publisher Full Text**

Sapède D, Cau E: **The pineal gland from development to function.** 1st ed, *Curr Top Dev Biol.* Elsevier Inc. 2013; **106**: 171–215.
PubMed Abstract | **Publisher Full Text**

Springer MS, Signore AV, Paijmans JLA, *et al.*: **Interordinal gene capture, the phylogenetic position of Steller's sea cow based on molecular and morphological data, and the macroevolutionary history of Sirenia.** *Mol Phylogenet Evol.* 2015; **91**: 178–193.
PubMed Abstract | **Publisher Full Text**

Springer MS, Starrett J, Morin PA, *et al.*: **Inactivation of *C4orf26* in toothless placental mammals.** *Mol Phylogenet Evol.* 2016; **95**: 34–45.
PubMed Abstract | **Publisher Full Text**

Stamatakis A: **RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies.** *Bioinformatics.* 2014; **30**(9): 1312–1313.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Stebbins RC, Eakin RM: **The role of the "third eye" in reptilian behavior.** *Am Museum Noitates.* 1958; 1–40.
Reference Source

Stefulj J, Hörtner M, Ghosh M, *et al.*: **Gene expression of the key enzymes of melatonin synthesis in extrapineal tissues of the rat.** *J Pineal Res.* 2001; **30**(4): 243–247.
PubMed Abstract | **Publisher Full Text**

St John JA, Braun EL, Isberg SR, *et al.*: **Sequencing three crocodilian genomes to illuminate the evolution of archosaurs and amniotes.** *Genome Biol.* 2012; **13**(1): 415.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Suofu Y, Li W, Jean-Alphonse FG, *et al.*: **Dual role of mitochondria in producing melatonin and driving GPCR signaling to block cytochrome c release.** *Proc Natl Acad Sci U S A.* 2017; **114**(38): E7997–E8006.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Tan DX, Hardeland R, Back K, *et al.*: **On the significance of an alternate pathway of melatonin synthesis via 5-methoxytryptamine: comparisons across species.** *J Pineal Res.* 2016; **61**(1): 27–40.
PubMed Abstract | **Publisher Full Text**

Tan DX, Zheng X, Kong J, *et al.*: **Fundamental issues related to the origin of melatonin and melatonin isomers during evolution: Relation to their biological functions.** *Int J Mol Sci.* 2014; **15**(9): 15858–15890.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Underwood H, Binkley S, Siopes T, *et al.*: **Melatonin rhythms in the eyes, pineal bodies, and blood of Japanese quail (*Coturnix coturnix japonica*).** *Gen Comp Endocrinol.* 1984; **56**(1): 70–81.
PubMed Abstract | **Publisher Full Text**

Valente R, Alves F, Sousa-Pinto I, *et al.*: **Functional or Vestigial? The Genomics of the Pineal Gland in Xenarthra.** *J Mol Evol.* 2021; **89**(8): 565–575.
PubMed Abstract | **Publisher Full Text**

Vonk FJ, Casewell NR, Henkel CV, *et al.*: **The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system.** *Proc Natl Acad Sci U S A.* 2013; **110**(51): 20651–20656.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Yang Z: **PAML 4: Phylogenetic analysis by maximum likelihood.** *Mol Biol Evol.* 2007; **24**(8): 1586–1591.
PubMed Abstract | **Publisher Full Text**

Yin D, Zhou R, Yin M, *et al.*: **Gene Duplication and Loss of AANAT in Mammals Driven by Rhythmic Adaptations.** *Mol Biol Evol.* 2021; **38**(9): 3925–3937.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Zhao D, Yu Y, Shen Y, *et al.*: **Melatonin synthesis and function: Evolutionary history in animals and plants.** *Front Endocrinol (Lausanne).* 2019; **10**: 249.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Zheng Y, Wiens JJ: **Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for squamate reptiles (lizards and snakes) based on 52 genes and 4162 species.** *Mol Phylogenet Evol.* 2016; **94**(Pt B): 537–547.
PubMed Abstract | **Publisher Full Text**

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 21 January 2022

https://doi.org/10.21956/openreseurope.15466.r28164

© **2022 Schott R.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Ryan K Schott**
Department of Biology, York University, York, ON, Canada

The authors have done a nice job addressing the concerns of myself and the other Reviewer. I am generally satisfied with the outcome and believe the paper to be improved.

While I find the PAML analyses easier to understand now, I still find the approach taken unusual. That is, using one alternate model and comparing it to numerous different null models rather than the other way around (or by successively comparing pairs of more complex models). I can better understand the reasoning behind this now, but there is a potential issue in that each null model employed by the authors has not been shown to itself be a significant improvement over a less complex model. Further, a general comparison among models (eg., using AIC values) was not performed and is not possible since the number of parameters in each model is not shown (as a side note: it would be nice to add these to the tables). With the tests done, there is often no way to distinguish between two 'null models' that place a branch in the background or fix it to w=1.

The authors indicated they were not interested in average patterns, but I think these could have been useful for the reader to get a broader picture of the patterns and to test for specific differences. In some cases, it is not clear why we would expect differences among lineages. For example, would we expect dn/ds to differ for pseudogenes among lineages?

That said I think the analysis performed by the authors is sufficient to support their conclusions and am happy to add my approval.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: molecular evolution, comparative genomics, sensory biology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 14 December 2021

https://doi.org/10.21956/openreseurope.15466.r28165

© **2021 Janiak M.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mareike C. Janiak
School of Science, Engineering and Environment, University of Salford, Salford, UK

I appreciate the authors' thorough responses to both reviews. I am satisfied with the changes made to the manuscript in response to my review and have no further comments.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: comparative genetics/genomics, dietary adaptations

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 20 July 2021

https://doi.org/10.21956/openreseurope.14874.r27206

© **2021 Schott R.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ryan K Schott
[1] Department of Biology, York University, York, ON, Canada
[2] Department of Biology, York University, York, ON, Canada

In this study, the authors examined two melatonin synthesis genes (*AANAT* and *ASMT*) and two melatonin receptors (*MTNR1A* and *MTNR1B*) in 110 placental mammals and 13 crocodylians. They hypothesized that species which lack pineal glands, the primary source of melatonin production, would lack functional melatonin synthesis and receptor genes. The authors found support for this hypothesis in placental mammals, but not in crocodylians. In the 13 crocodylians examined, all of which are thought to lack pineal glands, the authors found intact melatonin genes and expression of two of those genes was confirmed with tissue-specific transcriptome sequencing in American alligator, specifically in the eye. In contrast, placental mammals thought to lack pineal glands (xenarthrans, pangolins, sirenians, and whales), as well as colugos and some fossorial mammal lineages, showed evidence for independent inactivation and loss of all four melatonin genes.

Many other lineages also showed evidence of inactivation or loss in a smaller number of the genes. The authors speculate that loss of the melatonin synthesis and receptor genes may be associated with changes in thermoregulation and the loss of non-shivering thermogenesis. Overall, the manuscript is very well written, presents new and interesting results, and outlines several important areas for future research. I outline below several areas where the authors could further improve the paper:

- Inference of gene loss and pseudogenization depends on high quality sequencing data and genomic assemblies and apparent loss and pseudogenization could be due to incomplete sequencing/assembly or sequencing/assembly errors. While patterns of shared indels or premature stop codons are particularly convincing, I believe this caveat should still be addressed, especially considering that ~70% of the mammals examined were listed as having at least one pseudogenized or lost gene.

- Further to the point above, the sources of the genomic data used in the study is unclear. The main supplementary table of the genome sources does not provide much meaningful information (e.g., source listed as WGS). Looking through some of the subsequent SI tables accession numbers are given for some of the genes, but others are simply listed as New. Ideally all of the genomes would be made available, but minimally more detail needs to be provided on the genomic data (quality, completeness, etc...) used to obtain the sequences analyzed in the current study. It is difficult to evaluate the robustness of the results or reproduce the study without this.

- It was also unclear whether the transcriptome data were newly sequenced or the analysis was of previously published data. More detail should be provided.

- The explanation of the methodology and approach used in the selection analyses was difficult to follow and perhaps overly complex. The authors mention a nested approach but do not provide much detail. Looking through the supplementary tables of PAML results I do not see nested sets of models or an attempt to find the best fitting models. From what I can tell based on the tables, the approach relies on what appears to be a free ratio model, the use of which is generally discouraged due to it being overly parameter rich. A large number of likelihood ratio tests are then made between the free ratio model and models placing individual branches or clades in the 'foreground'. This approach appears to (incorrectly) use the most parameter-rich model as the null model for the likelihood ratio tests. Models that restrict the 'foreground' branch/clade to $\omega = 1$ are also used but they are also compared to the parameter rich model (rather than to same model but where the foreground ω is free to vary). As far as I can tell, no models are used that combine multiple branches or clades as might be expected. Beyond these issues, a much simpler approach could be used that also more directly tests the author's hypotheses by comparing ω between mammals with and without pineal glands and between lineages with and without evidence for pseudogenization in that specific gene. A codon model specifically designed to detect relaxation of selective constraint (ie, RELAX) would be useful in this regard.

- It would be nice to see selection analyses on the crocodilian sequences as well to, for example, show that they are still under strong selective constraint. I think this could still be useful without other sequences for comparison (eg, from birds/turtles).

- I initially found the mix of analyses on mammalian and crocodilian genes unusual, but this

made more sense within the context of animals with and without pineal glands. This could perhaps be made more clear, especially in the title which only mentions placental mammals and not crocodylians or the lack of pineal glands.

- I generally thought the figures were well put together quite effective, but the text size (most notably in Fig. 2) can be quite small and hard to read, even when zooming in.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: molecular evolution, comparative genomics, sensory biology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 24 Nov 2021

Christopher Emerling

Author responses to Dr. Schott in bold text below with reviewer comments in italics.

In this study, the authors examined two melatonin synthesis genes (AANAT and ASMT) and two melatonin receptors (MTNR1A and MTNR1B) in 110 placental mammals and 13 crocodylians. They hypothesized that species which lack pineal glands, the primary source of melatonin production, would lack functional melatonin synthesis and receptor genes. The authors found support for this hypothesis in placental mammals, but not in crocodylians. In the 13 crocodylians examined, all of which are thought to lack pineal glands, the authors found intact melatonin genes and expression of two of those genes was confirmed with tissue-specific transcriptome sequencing in American alligator, specifically in the eye. In contrast, placental mammals thought

to lack pineal glands (xenarthrans, pangolins, sirenians, and whales), as well as colugos and some fossorial mammal lineages, showed evidence for independent inactivation and loss of all four melatonin genes. Many other lineages also showed evidence of inactivation or loss in a smaller number of the genes. The authors speculate that loss of the melatonin synthesis and receptor genes may be associated with changes in thermoregulation and the loss of non-shivering thermogenesis. Overall, the manuscript is very well written, presents new and interesting results, and outlines several important areas for future research. I outline below several areas where the authors could further improve the paper:

We extend our thanks to Dr. Schott for spending the time to review our manuscript and giving generally encouraging feedback on our presentation. We appreciate his critical feedback and have done our best to appropriately address all of his comments below.

Inference of gene loss and pseudogenization depends on high quality sequencing data and genomic assemblies and apparent loss and pseudogenization could be due to incomplete sequencing/assembly or sequencing/assembly errors. While patterns of shared indels or premature stop codons are particularly convincing, I believe this caveat should still be addressed, especially considering that ~70% of the mammals examined were listed as having at least one pseudogenized or lost gene.

We agree that this can indeed be an issue with genomic data. While we are confident that the shared inactivating mutations, often paired with dN/dS estimates suggestive of relaxed selection, are strong evidence of pseudogenization in the pineal-less mammals and colugos, we believe Dr. Schott is correct that we should present additional evidence for the inactivation of genes in other taxa. To address this, we have done the following:

1. **Collected additional genomic sequences from related taxa for species with seemingly unique inactivating mutations (i.e., to test for evidence of shared mutations).**

2. **In cases where this did not provide positive evidence of inactivating mutations (e.g., no adequate sister taxon, no shared mutations), we mapped short reads to the genomic contigs for such taxa.**

We performed these analyses for all taxa that have evidence for inactivation of one or more melatonin synthesis gene and/or both melatonin receptor genes, given that these are the most relevant for our study. This is in contrast to the myriad examples of species that only have an *MTNR1A* or *MTNR1B* pseudogene, which would likely still have some, if not most, melatonin pathway functionality We have included these additional data in the Underlying Data and Extended Data in the form of Supplementary Table S13 and Supplementary Figures S8–S14. We believe our analyses effectively corroborate the inactivation of nearly every example of melatonin synthesis and/or signalling disruption, with the sole exception being for the pair of pseudogenic melatonin receptor genes in the Hawaiian monk seal *Neomonachus schauinslandi* (notably an unexpected finding).

Further to the point above, the sources of the genomic data used in the study is unclear. The

main supplementary table of the genome sources does not provide much meaningful information (e.g., source listed as WGS). Looking through some of the subsequent SI tables accession numbers are given for some of the genes, but others are simply listed as New. Ideally all of the genomes would be made available, but minimally more detail needs to be provided on the genomic data (quality, completeness, etc…) used to obtain the sequences analyzed in the current study. It is difficult to evaluate the robustness of the results or reproduce the study without this.

We agree that this would be of utility to some readers. To address this, we have added more information to Supplementary Table S1, including the identifying prefix for assemblies found in NCBI, coverage, contig N50 and scaffold N50. For sequences assembled via short read mapping, we have included coverage data for each exon that we were able to assemble from short reads. Furthermore, since our novel sequences were derived from different genomic projects in different labs that have yet to publish their assemblies, we have indicated in the Supplementary Table S1 legend whom to contact (with e-mail addresses) for each relevant sequencing project.

It was also unclear whether the transcriptome data were newly sequenced or the analysis was of previously published data. More detail should be provided.

The transcriptome data for *A. mississippiensis* were derived from NCBI's sequence read archive (SRA). Though this was described in the Underlying Data, we now have further clarified this in the revised manuscript in the following ways:

- **A statement in the Materials and Methods on the generation of crocodylian sequences (on pg. 7).**

- **An explicit statement that the RNA sequencing data for the Alligator has been previously published with a reference to the corresponding paper (pg. 10).**

The explanation of the methodology and approach used in the selection analyses was difficult to follow and perhaps overly complex. The authors mention a nested approach but do not provide much detail. Looking through the supplementary tables of PAML results I do not see nested sets of models or an attempt to find the best fitting models. From what I can tell based on the tables, the approach relies on what appears to be a free ratio model, the use of which is generally discouraged due to it being overly parameter rich. A large number of likelihood ratio tests are then made between the free ratio model and models placing individual branches or clades in the 'foreground'. This approach appears to (incorrectly) use the most parameter-rich model as the null model for the likelihood ratio tests. Models that restrict the 'foreground' branch/clade to ω = 1 are also used but they are also compared to the parameter rich model (rather than to same model but where the foreground ω is free to vary). As far as I can tell, no models are used that combine multiple branches or clades as might be expected. Beyond these issues, a much simpler approach could be used that also more directly tests the author's hypotheses by comparing ω between mammals with and without pineal glands and between lineages with and without evidence for pseudogenization in that specific gene. A codon model specifically designed to detect relaxation of selective constraint (ie, RELAX) would be useful in this regard.

We thank Dr. Schott for this critical feedback, as Dr. Janiak had the same concerns; it clearly needed to be clarified further. In short, we did not use a free ratio model as the null model. We designated a model in which certain branches were grouped based on their association with distinct inferences of pineal gland loss / gene losses, with

branches being defined as functional (no evidence of dysfunction), pseudogenic (post-dating branch in which pseudogenization is inferred to have occurred), mixed/transitional (branch in which pseudogenization is inferred to have occurred), and pre-mutation (no evidence of pseudogenization on the branch but external evidence suggests gene may be under relaxed selection). As stated in the manuscript, this is following the approach of Meredith *et al.* 2009. Accordingly, there was a 'master model' for each gene, with each branch category described above, and this model was then compared to null models in which a specific pseudogene branch (or set of branches), mixed/transitional branch or pre-mutation branch was set to the background or 1, effectively testing whether a specific branch or set of branches was distinct from the background and/or 1, and therefore displaying evidence of relaxed selection. This was to test for individual historical cases of relaxed selection, rather than a broad 'average' pattern of relaxed selection, the latter of which does not accomplish what we were trying to test. Details of what we added can be found in the sections described below:

1. **We have elaborated on the textual specifics of these methods under a section entitled "Evolutionary analyses" within the Materials and Methods (pgs. 9–10).**

2. **We have included four supplementary figures showing how each of the branch categories within the dN/dS analyses were designated within the 'master' model for each gene (Supplementary Figures S1–S4). We then paired the branch labeling with the dN/dS results in Supplementary Tables S7–S10 by use of numbering, to make for a smoother comparison of the two.**

We believe that this should be far easier to understand than in our previous iteration, and hope it will satisfy Drs. Schott and Janiak.

It would be nice to see selection analyses on the crocodilian sequences as well to, for example, show that they are still under strong selective constraint. I think this could still be useful without other sequences for comparison (eg, from birds/turtles).

We agree with this suggestion and have done so. We added sections in the Materials and Methods (pgs 8 & 10) and Results (pg 13), added the alignments and topology to Supplementary Datasets S5–S9, and modelling results in Supplementary Tables S6 and S11. The basic result is that all four genes appear to be under purifying selection in crocodylians, with three ratio models (background, stem Crocodylia, crown Crocodylia) being statistically indistinguishable from one ratio models, and the crown and stem Crocodylia w estimates being very similar to the background w.

I initially found the mix of analyses on mammalian and crocodilian genes unusual, but this made more sense within the context of animals with and without pineal glands. This could perhaps be made more clear, especially in the title which only mentions placental mammals and not crocodylians or the lack of pineal glands.

Regarding the title, given that the first version of the manuscript is technically published via the Open Research Europe model, we thought changing it would make it confusing for future referencing purposes, especially as the manuscript has already been cited. Furthermore, given that the crocodylians we examined did not show

evidence of inactivating melatonin pathway genes, including a crocodylian reference in the title seemed to us to take away from our most important finding, i.e., mentioning the absence of positive evidence for inactivation in crocs may detract from the title. However, we have tried to make the inclusion of crocodylians in our study clearer in the following ways:

1. **Including an initial mention of the crocodylian results at the end of the introduction (pg 6).**

2. **Including a paragraph in the Materials and Methods describing the sequencing strategy for our novel crocodylian sequences (pg 8).**

3. **As discussed above, there are now paragraphs discussing the dN/dS analyses unique to crocs and other sauropsids in the Materials and Methods (pgs 8 & 10).**

4. **Including a section devoted to crocs in the Results (pg 13).**

I generally thought the figures were well put together quite effective, but the text size (most notably in Fig. 2) can be quite small and hard to read, even when zooming in.

We are happy to hear that Dr. Schott found the figures to be effective. We have fixed the issues with text size as follows

1. **We made the letter labels in Figure 1 larger and superimposed colored circles upon to make them stand out among the branches.**

2. **For Figure 2, we made species names larger and increased the resolution for the alignments and species names.**

Competing Interests: No competing interests were disclosed.

Reviewer Report 16 July 2021

https://doi.org/10.21956/openreseurope.14874.r27204

© 2021 Janiak M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Mareike C. Janiak
[1] School of Science, Engineering and Environment, University of Salford, Salford, UK
[2] School of Science, Engineering and Environment, University of Salford, Salford, UK

In the manuscript “Genomic evidence for the parallel regression of melatonin synthesis and signaling pathways in placental mammals” the authors examine four genes related to melatonin synthesis and signaling in a broad sample of placental mammals, as well as 13 crocodylians. The authors pay special attention to species that (appear to) lack a pineal gland, the organ that

secretes melatonin. Given the importance of melatonin in regulating an organism's circadian rhythm, among other function, the authors hypothesized that the four genes involved in melatonin synthesis and signaling may remain functional even in species that lack a pineal gland, as there may be other sources of melatonin within the body.

The authors show that, despite lacking a pineal gland, crocodylians show no disruptions to their melatonin synthesis and signaling genes and may synthesize melatonin within the eye. Several groups of placental mammals, however, appear to have disrupted melatonin synthesis and/or signaling pathways, including those lacking a pineal gland, living underground, and colugos.

The authors discuss the implications of these findings and suggest directions for future research, such as potential links with temperature and thermoregulation. However, the authors also rightly point out that these results are surprising, considering the (assumed) importance of melatonin signaling and suggest that there may be other mechanisms for melatonin detection, or other pathways for circadian signaling.

Overall, I found the manuscript to be very well-written and clearly presented, making it quite enjoyable to read. The figures are also very effective in communicating the results visually.

My main suggestions for improving the manuscript are as follows:

1. The authors may be interested in the program RELAX (Wertheim et al., 2015; https://doi.org/10.1093/molbev/msu400) to explicitly test for relaxed selection on the pineal gland-less lineages. It seems like a useful test for the present dataset.

2. I found the description of the PAML analyses confusing and I'm not sure that I would be able to reproduce them based on the methods, specifically the paragraph describing the nested models and models for relaxed selection (note to the editor: it would be very useful if the manuscripts provided to reviewers included line numbers!). Also, it sounds like these included a large number of comparisons, were these results corrected for multiple comparisons? And if so, how?

3. The authors refer to all genes that show premature stop codons as "pseudogenes" even though this may not be what they are. There is growing evidence that so-called "pseudogenes" may in fact be functional, for example see Goodhead & Darby, 2015; https://doi.org/10.1016/j.mib.2014.11.012 and Cheetham et al., 2020; https://doi.org/10.1038/s41576-019-0196-1. While the authors are probably correct that these pseudogenes may no longer produce an equivalent product, I think it's worth acknowledging the functional potential of pseudogenes.

Some minor suggestions:

1. I am a bit confused about the source of the crocodylian sequences, especially the transcriptomes. Based on the methods, it sounds as if these were sequenced by other researchers and retrieved from NCBI by the authors. However, the acknowledgements include thanks to a zoological park for crocodylian tissue samples. Are any of the results based on de novo sequencing? If so, it would be appropriate to describe this in detail in the methods.

2. The letters along the branches in Figure 1 are quite small and hard to read without zooming in. Could these be a bit bigger?

3. Finally, I am admittedly not an expert in melatonin pathway or circadian rhythm research, so I hope that others can provide more insightful feedback on the authors' discussion of the implications of their findings. In my non-expert opinion the authors appear appropriately cautious in how they interpret their findings, posing more questions than answers.

References

1. Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, et al.: RELAX: detecting relaxed selection in a phylogenetic framework.*Mol Biol Evol*. 2015; **32** (3): 820-32 PubMed Abstract | Publisher Full Text
2. Goodhead I, Darby AC: Taking the pseudo out of pseudogenes.*Curr Opin Microbiol*. 2015; **23**: 102-9 PubMed Abstract | Publisher Full Text
3. Cheetham SW, Faulkner GJ, Dinger ME: Overcoming challenges and dogmas to understand the functions of pseudogenes.*Nat Rev Genet*. **21** (3): 191-201 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: comparative genetics/genomics, dietary adaptations

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 24 Nov 2021

Christopher Emerling

AUTHOR RESPONSES TO DR. JANIAK IN BOLD TEXT BELOW Author responses to Dr.

Janiak in bold text below with reviewer comments in italics.

In the manuscript "Genomic evidence for the parallel regression of melatonin synthesis and signaling pathways in placental mammals" the authors examine four genes related to melatonin synthesis and signaling in a broad sample of placental mammals, as well as 13 crocodylians. The authors pay special attention to species that (appear to) lack a pineal gland, the organ that secretes melatonin. Given the importance of melatonin in regulating an organism's circadian rhythm, among other function, the authors hypothesized that the four genes involved in melatonin synthesis and signaling may remain functional even in species that lack a pineal gland, as there may be other sources of melatonin within the body.

The authors show that, despite lacking a pineal gland, crocodylians show no disruptions to their melatonin synthesis and signaling genes and may synthesize melatonin within the eye. Several groups of placental mammals, however, appear to have disrupted melatonin synthesis and/or signaling pathways, including those lacking a pineal gland, living underground, and colugos.

The authors discuss the implications of these findings and suggest directions for future research, such as potential links with temperature and thermoregulation. However, the authors also rightly point out that these results are surprising, considering the (assumed) importance of melatonin signaling and suggest that there may be other mechanisms for melatonin detection, or other pathways for circadian signaling.

Overall, I found the manuscript to be very well-written and clearly presented, making it quite enjoyable to read. The figures are also very effective in communicating the results visually.

We thank Dr. Janiak for taking the time to review our manuscript and giving positive feedback on its overall presentation. We have done our best to adequately respond to her insightful comments below.

My main suggestions for improving the manuscript are as follows:

1. *The authors may be interested in the program RELAX (Wertheim et al., 2015; https://doi.org/10.1093/molbev/msu400) to explicitly test for relaxed selection on the pineal gland-less lineages. It seems like a useful test for the present dataset.*

We are indeed familiar with RELAX and have found it to be useful in certain analyses. However, we believe that PAML is sufficient in the context of this particular study in providing additional evidence of pseudogenization besides disrupting mutations. Detailed comparisons of PAML and RELAX within the context of pseudogene evolution to better understand how the analyses compare in estimating dN/dS during transitions from functional to nonfunctional genes would indeed be interesting to conduct. However, we think that this is beyond the scope of this empirical melatonin-focused paper.

I found the description of the PAML analyses confusing and I'm not sure that I would be able to reproduce them based on the methods, specifically the paragraph describing the nested models and models for relaxed selection (note to the editor: it would be very useful if the manuscripts provided to reviewers included line numbers!). Also, it sounds like these included a large number of comparisons, were these results corrected for multiple comparisons? And if so, how

We agree with this fair criticism on the point of the methods. Given that Dr. Schott made the same point, clearly this needed to be fixed to ensure replicability. We have made a concerted effort to improve comprehension of the methods in the following ways:

1. **We have elaborated on the textual specifics of these methods under a section entitled "Evolutionary analyses" within the Materials and Methods (pgs. 9–10).**

2. **We have included four supplementary figures showing how each of the branch categories within the dN/dS analyses were designated within the 'master' model for each gene (Supplementary Figures S1–S4). We then paired the branch labeling with the dN/dS results in Supplementary Tables S7–S10 by use of numbering, to make for a smoother comparison of the two.**

Regarding multiple test comparisons, we have applied the Holm-Bonferroni method for multiple testing. This is now detailed in the Materials and Methods (pg. 10), Results (pgs. 14 & 15) and the adjusted p-values are now included in Supplementary Tables S7–S10.

The authors refer to all genes that show premature stop codons as "pseudogenes" even though this may not be what they are. There is growing evidence that so-called "pseudogenes" may in fact be functional, for example see Goodhead & Darby, 2015; https://doi.org/10.1016/j.mib.2014.11.012 and Cheetham et al., 2020; https://doi.org/10.1038/s41576-019-0196-1. While the authors are probably correct that these pseudogenes may no longer produce an equivalent product, I think it's worth acknowledging the functional potential of pseudogenes.

Thank you for this suggestion. Given that Cheetham *et al.* 2020 is more appropriate for our paper, we have included a reference to this study in the discussion (pg. 15), reproduced below: "Despite this, given that recent discoveries have shown that presenting as a pseudogene does not always indicate that all biological function is lost (Cheetham *et al*., 2020), there remains a possibility that these apparently dysfunctional genes are able to contribute to melatonin metabolism in some unknown fashion."

I am a bit confused about the source of the crocodylian sequences, especially the transcriptomes. Based on the methods, it sounds as if these were sequenced by other researchers and retrieved from NCBI by the authors. However, the acknowledgements include thanks to a zoological park for crocodylian tissue samples. Are any of the results based on de novo sequencing? If so, it would be appropriate to describe this in detail in the methods.

Most of the crocodylian genomic sequences were generated by us, with the exception of four draft genome assemblies (*Alligator mississippiensis, A. sinensis, Gavialis gangeticus, Crocodylus porosus*), hence the acknowledgement of the St. Augustine Alligator Farm Zoological Park for tissue samples. The transcriptome data for *A. mississippiensis* were derived from NCBI's sequence read archive (SRA). Though this was described in the Underlying Data, we now have further clarified this in the revised manuscript in the following ways:

1. **A statement in the Materials and Methods on the generation of crocodylian sequences (on pg. 7).**

2. **An explicit statement that the RNA sequencing data for the Alligator has been previously published with a reference to the corresponding paper (pg. 10).**

The letters along the branches in Figure 1 are quite small and hard to read without zooming in. Could these be a bit bigger?

We have addressed this by making the letters larger and superimposed colored circles upon to make them stand out among the branches.

Finally, I am admittedly not an expert in melatonin pathway or circadian rhythm research, so I hope that others can provide more insightful feedback on the authors' discussion of the implications of their findings. In my non-expert opinion the authors appear appropriately cautious in how they interpret their findings, posing more questions than answers.

We agree that this would have been ideal. We requested reviews from three different melatonin experts, but unfortunately none were available. Due to time constraints, we have decided to revise the manuscript with the reviews provided so far and will look forward to commentary from future studies.

Competing Interests: No competing interests were disclosed.