

Eye-Transcriptome and Genome-Wide Sequencing for Scolecophidia: Implications for Inferring the Visual System of the Ancestral Snake

David J. Gower ^{1,*}, James F. Fleming^{2,3}, Davide Pisani ^{2,4}, Freek J. Vonk⁵, Harald M. I. Kerckamp⁶, Leo Peichl ^{7,8}, Sonja Meimann⁷, Nicholas R. Casewell ⁹, Christiaan V. Henkel ^{6,10}, Michael K. Richardson ⁶, Kate L. Sanders¹¹, and Bruno F. Simões ^{2,11,12,*}

¹Life Sciences, The Natural History Museum, London, United Kingdom

²School of Life Sciences, University of Bristol, Bristol, United Kingdom

³Institute for Advanced Biosciences, Keio University, Yamagata, Japan

⁴School of Earth Sciences, University of Bristol, Bristol, United Kingdom

⁵Naturalis Biodiversity Center, Leiden, The Netherlands

⁶Institute of Biology, University of Leiden, Leiden, The Netherlands

⁷Institute of Cellular and Molecular Anatomy, Dr. Senckenberg Anatomy, Goethe University Frankfurt, Frankfurt am Main, Germany

⁸Institute of Clinical Neuroanatomy, Dr. Senckenberg Anatomy, Goethe University Frankfurt, Frankfurt am Main, Germany

⁹Centre for Snakebite Research & Interventions, Liverpool School of Tropical Medicine, Liverpool, United Kingdom

¹⁰Faculty of Veterinary Medicine, Norwegian University of Life Sciences, Ås, Norway

¹¹School of Biological Sciences, University of Adelaide, Adelaide, South Australia, Australia

¹²School of Biological and Marine Sciences, University of Plymouth, Plymouth, United Kingdom

*Corresponding authors. Email: d.gower@nhm.ac.uk; bruno.simoies@plymouth.ac.uk.

Accepted: 8 October 2021

Abstract

Molecular genetic data have recently been incorporated in attempts to reconstruct the ecology of the ancestral snake, though this has been limited by a paucity of data for one of the two main extant snake taxa, the highly fossorial Scolecophidia. Here we present and analyze vision genes from the first eye-transcriptomic and genome-wide data for Scolecophidia, for *Anilius bicolor*, and *A. bituberculatus*, respectively. We also present immunohistochemistry data for retinal anatomy and visual opsin-gene expression in *Anilius*. Analyzed in the context of 19 lepidosaurian genomes and 12 eye transcriptomes, the new genome-wide and transcriptomic data provide evidence for a much more reduced visual system in *Anilius* than in non-scolecophidian (=alethinophidian) snakes and in lizards. In *Anilius*, there is no evidence of the presence of 7 of the 12 genes associated with alethinophidian photopic (cone) phototransduction. This indicates extensive gene loss and many of these candidate gene losses occur also in highly fossorial mammals with reduced vision. Although recent phylogenetic studies have found evidence for scolecophidian paraphyly, the loss in *Anilius* of visual genes that are present in alethinophidians implies that the ancestral snake had a better-developed visual system than is known for any extant scolecophidian.

Key words: gene loss, opsins, phylogeny, regressive evolution, Squamata, vision.

Introduction

The origin of major lineages (higher taxa) is a topic of broad interest in evolutionary biology (e.g., Darwin 1859; Smith and

Szathmary 1997; Holland et al. 2008; Kemp 2015; Lane 2015; Eme et al. 2017), and the origin of snakes (Serpentes) is exemplary in this respect (e.g., Bellairs and Underwood

© The Author(s) 2021. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Significance

The origin of snakes is subject to ongoing, high-profile debate. We present the first eye-transcriptomic and substantial genome-wide sequencing data for any scolecophidian snake (Scolecophidia are dedicated burrowers and form one-half of the oldest divergence in the snake tree). Comparative analyses of the presence of functional vision genes among lizards and snakes strongly indicate that the visual system (especially that adapted for bright-light vision) of scolecophidians is too reduced to be a good model for the visual system of the ancestral snake, thus challenging the hypothesis that snakes evolved from extreme burrowers.

1951; Bellairs 1972; Rieppel 1988; Caprette et al. 2004; Lee 2005; Caldwell 2007; Hsiang et al. 2015; Simões et al. 2015; Yi and Norell 2015; Emerling 2017; da Silva et al. 2018; Miralles et al. 2018). Research into snake origins has a long history (Rieppel 1988) with several recent contributions, especially to debates about the likely eco(morpho)logical attributes of the most-recent common ancestor of extant snakes (the “ancestral snake” hereafter). Most recent research into the ecological aspect of snake origins has focused on trying to infer ancestral states based on the anatomy of extant and extinct snakes and their putative closest relatives (e.g., Hsiang et al. 2015; Yi and Norell 2015; Harrington and Reeder 2017; da Silva et al. 2018; Miralles et al. 2018). Thus far, there has been relatively little input from molecular genetics into reconstructions of the ancestral snake. There is evidence that two of the ancestral five vertebrate (and squamate) visual opsin genes have been lost in the snake stem (Davies et al. 2009; Castoe et al. 2013; Simões et al. 2015, 2016a; Schott et al. 2018) and evidence of inactivation or loss of some claw-specific keratin genes, some gustatory (taste-related) genes, and some light-associated genes in some extant snakes (Emerling 2017).

The earliest phylogenetic divergence among extant snakes is between Alethinophidia and Scolecophidia, or between the two major lineages of scolecophidians given that most recent molecular genetic evidence supports their paraphyly—see below. Scolecophidians comprise approximately 460 of the approximately 3,850 currently recognized extant snake species. They are typically small, cylindrical, fossorial (=burrowing) snakes with reduced eyes and small mouths, strikingly different in ecology and morphology from the more familiar alethinophidians, which comprise all other crown snakes including, for example, pythons, vipers, and cobras (e.g., Miralles et al. 2018). One of the major limitations in reconstructions of the ancestral snake from molecular genetic data (Emerling 2017) or in testing for signals of divergent molecular genetic selection along the snake phylogenetic stem (Schott et al. 2018) has been the lack of genomic data for scolecophidian snakes.

A small but increasing sampling of genomic data (Castoe et al. 2013; Perry et al. 2018; Schott et al. 2018; Emerling 2017) presents evidence for differences (including gene absences) between snakes and nonsnake squamates (lizards).

However, thus far these data are available only for a relatively small sample of alethinophidians, such that ascertaining whether inferred genomic changes occurred in the snake stem or the alethinophidian stem (within the snake crown group) has not been possible. Furthermore, interpretation of the few molecular genetic data that are available has been complicated by disagreement among phylogeneticists as to whether Scolecophidia is a monophyletic (i.e., comprising one half of the basal divergence among crown snakes: e.g., Lee et al. 2007) or paraphyletic (e.g., Wiens et al. 2012; Miralles et al. 2018; Burbrink et al. 2020) outgroup of all other crown snakes (Alethinophidia). Thus, the extent to which phenotypes of extant scolecophidians might represent highly derived adaptations to fossoriality and/or potentially plesiomorphic conditions present in the ancestral snake remains contentious.

Here, we contribute to the debate by presenting analyses of vision-associated gene sequence data from the first eye-transcriptome and genome-wide sequence data and retinal immunohistochemistry data for Scolecophidia. The visual system has played a prominent role in debates about snake origins (Bellairs and Underwood 1951; Caprette et al. 2004; Simões et al. 2015; Miralles et al. 2018), and we expect genomic data on vision genes to provide insight into the ecophenotype of the ancestral snake. For example, ancestral fossoriality can be predicted to have resulted in the last common ancestor of extant snakes having a proportionately greater loss of functionality in genes associated with bright-light (photopic) than dim-light (scotopic) vision, such as occurs, for example, in highly fossorial mammals (Emerling and Springer 2014).

Results

Vision-Gene Complements

The presence or absence results for vision genes based on BLAST and phylogenetic analyses are summarized in [table 1](#). Of the 48 vision-associated genes studied here, all but seven occur in snakes. These seven are the genes coding for the phosphodiesterase subunit PDE6A, the rhodopsin kinase GRK1, the transducin gamma subunit GNGT1, the solute carrier SLC24A1, the retinol dehydrogenase RDH11, and the

visual opsins RH2 and SWS2. Four of these genes are absent also in lizards, *pde6a*, *gnpt1*, *slc24a1*, and *rdh11*. Among the 41 vision-associated genes occurring in at least some of the sampled snake genomes and eye transcriptomes, we found evidence for 28 in the *Anilius bicolor* eye-transcriptome and 29 in the *A. bituberculatus* genome-wide data (table 1; 27 of these genes common to both). One visual-cycle gene (*rbp1*) and one cone and rod phototransduction gene (*gucy2f*) found in the *A. bituberculatus* genome-wide sequencing were not found in the *A. bicolor* eye transcriptome, and so are perhaps not being expressed in the *Anilius* eye. Similarly, *gucy2f* is found in the genome but not eye transcriptome of *Thamnophis sirtalis* (table 1). Despite the relatively low coverage of the genome-wide data for *A. bituberculatus*, only one visual-cycle gene (*Irat*) was not detected that was however found in the *A. bicolor* eye transcriptome.

We inspected sequences of some of the 29 (selected at random) genes in the *A. bituberculatus* genome-wide data (supplementary table S3, Supplementary Material online) for evidence consistent with sequencing and/or assembly error. We took vision-gene sequences from the *A. bicolor* transcriptome and used blastn to search for the corresponding genes in the *A. bituberculatus* genome-wide data. In all examined cases (two examples are shown in supplementary appendix S4, Supplementary Material online), observed fragmentation in genome-wide sequences can be explained (beyond low coverage) fully by intron–exon boundaries, and coverage for the ends of fragments is typical of overall coverage. Additionally, we extracted the best-possible homologous sequences of the leading candidates for possible *Anilius* vision pseudogenes—the two genes in the *A. bituberculatus* genome-wide data that are absent in the *A. bicolor* transcriptome (*rbp1* and *gucy2f*)—via exon-by-exon tblastx queries of the genome-wide data using sequences from the closest available relatives (*Python* and *Boa*). The extracted and assembled *A. bituberculatus* *rbp1* and *gucy2f* sequences lack unexpected stop codons or frameshift indels, and phylogenetic analysis confirms their identity and demonstrates that they do not form unexpectedly long terminal phylogenetic branches (supplementary fig. S1, Supplementary Material online) that would be indications of pseudogenes or sequencing error. Among the genes for which we found evidence in *Anilius* spp. is the visual opsin gene *lws*, previously thought to be absent in scolecophidians (Simões et al. 2015). Phylogenetic trees (supplementary figs. S2–S4, Supplementary Material online) for snake visual opsins confirm the identity of the *A. bicolor* and *A. bituberculatus* *rh1* and *lws*.

There are 12 vision genes for which we found no evidence in *Anilius* spp. but which are present in at least one other, non-scolecophidian snake. None of these 12 is associated exclusively with the rod phototransduction pathway (of eight such genes in total), seven are associated with cone phototransduction (of 12 total, 58.3%), two with both rod and cone phototransduction (of 10 total, 20%), and three (of 12 total, 25%) with the visual cycle (table 1; supplementary tables S4 and S5,

Supplementary Material online). Of the 20 exclusively rod or cone phototransduction genes found in snakes, the probability of the seven gene absences in *Anilius* spp. all being cone phototransduction genes by chance is $0.01 (13/20 \times 12/19 \times 11/18 \times 10/17 \times 9/16 \times 8/15 \times 7/14 \times 6/13)$, such that it is extremely unlikely that the apparent diminution of cone phototransduction genes in the genomes of this genus is explained by random failure to detect cone genes via eye transcriptomes or relatively low-coverage genome-wide sequencing.

Inspection of the retinal protein networks for phototransduction (fig. 1) highlights the proportionately greater absence of evidence for genes more typically associated with photopic (typically cone-based) vision in *Anilius* spp. For both phototransduction and visual-cycle networks, there is a substantial degree of overlap between genes likely lost in *Anilius* and those seemingly lost in the genomes of the highly fossorial mammals possessing reduced visual systems studied by Emerling and Springer (2014).

Retinal Immunohistochemistry

The two sampled *Anilius* species provided similar results. General eye anatomical information, including size, is presented in supplementary appendix S3, Supplementary Material online. There are no obvious differences among outer nuclear layer (ONL) nuclei or notable heterogeneity in size and appearance of inner segments (supplementary fig. S5c, Supplementary Material online), that is, there are no indications for a dichotomy with morphologically different rods and cones comprising a duplex retina. Immunolabeling revealed a dense population of RH1-containing photoreceptors (fig. 2a and b) and a less dense but still substantial population of LWS-containing photoreceptors (fig. 2c and d; supplementary fig. S5a, Supplementary Material online). The RH1 (rod) opsin antiserum labeled outer segments most intensely, but also labeled many somata more faintly, a feature seen in many vertebrates including other snakes (Gower et al. 2019). The SWS1 opsin antiserum showed no labeling of photoreceptors (or other retinal cells) in sections and the wholemount (supplementary fig. S5a, Supplementary Material online). Immunohistochemistry corroborates the genetic analysis in terms of the absence of functional *sws1*, and the presence of functional *rh1* and *lws*, in both species.

Many photoreceptors were labeled for both RH1 and LWS, indicating co-expression of rod and cone opsin (fig. 2e1–e3; supplementary fig. S5d1–d3, Supplementary Material online). Some photoreceptors were exclusively labeled by either the RH1 antibody or LWS antiserum (colored arrowheads in fig. 2e3). It was not feasible to assess population densities of RH1 and LWS photoreceptors, nor to morphologically identify the photoreceptors as cones or rods, or to assess whether double cones or double rods are present. The LWS-labeled outer segments are long and thin, superficially more rod- than cone-like (supplementary fig. S5a, Supplementary Material

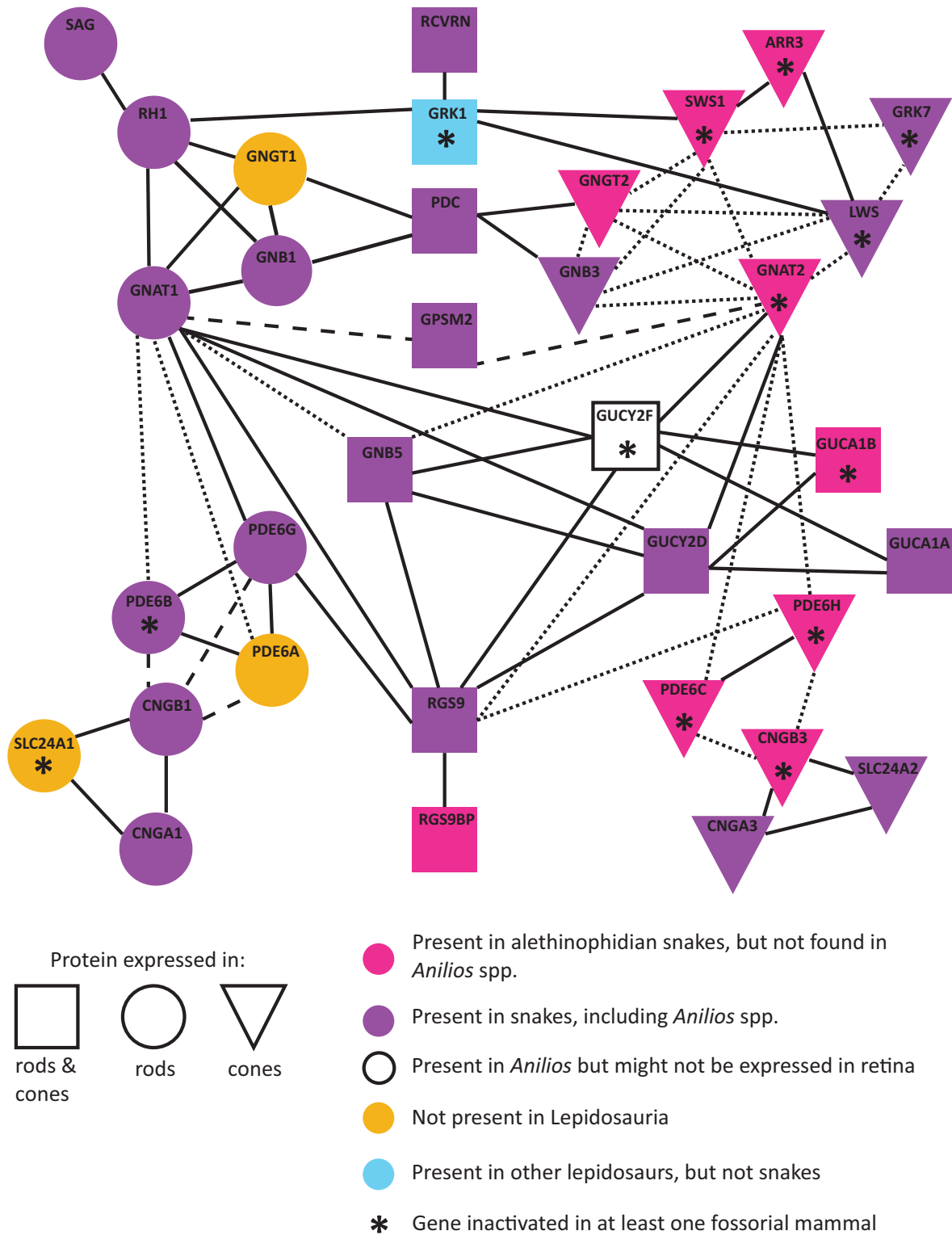


Fig. 1.—Network (modified from Emerling and Springer 2014: fig. 1) for 34 phototransduction proteins indicating absences in squamates and fossorial mammals as determined from genomic data. Note that loss of PDE6B and GRK1 in fossorial mammals is provisional, based on negative BLAST results rather than synteny analyses (Emerling and Springer 2014). Of the 36 phototransduction genes in table 1, protein products for the cone phototransduction (visual opsins) genes *sws2* and *rh2* are not shown because they are absent in all snakes and mammals.

Downloaded from https://academic.oup.com/gbe/article/13/12/evab253/6430116 by Liverpool School of Tropical Medicine user on 07 December 2021

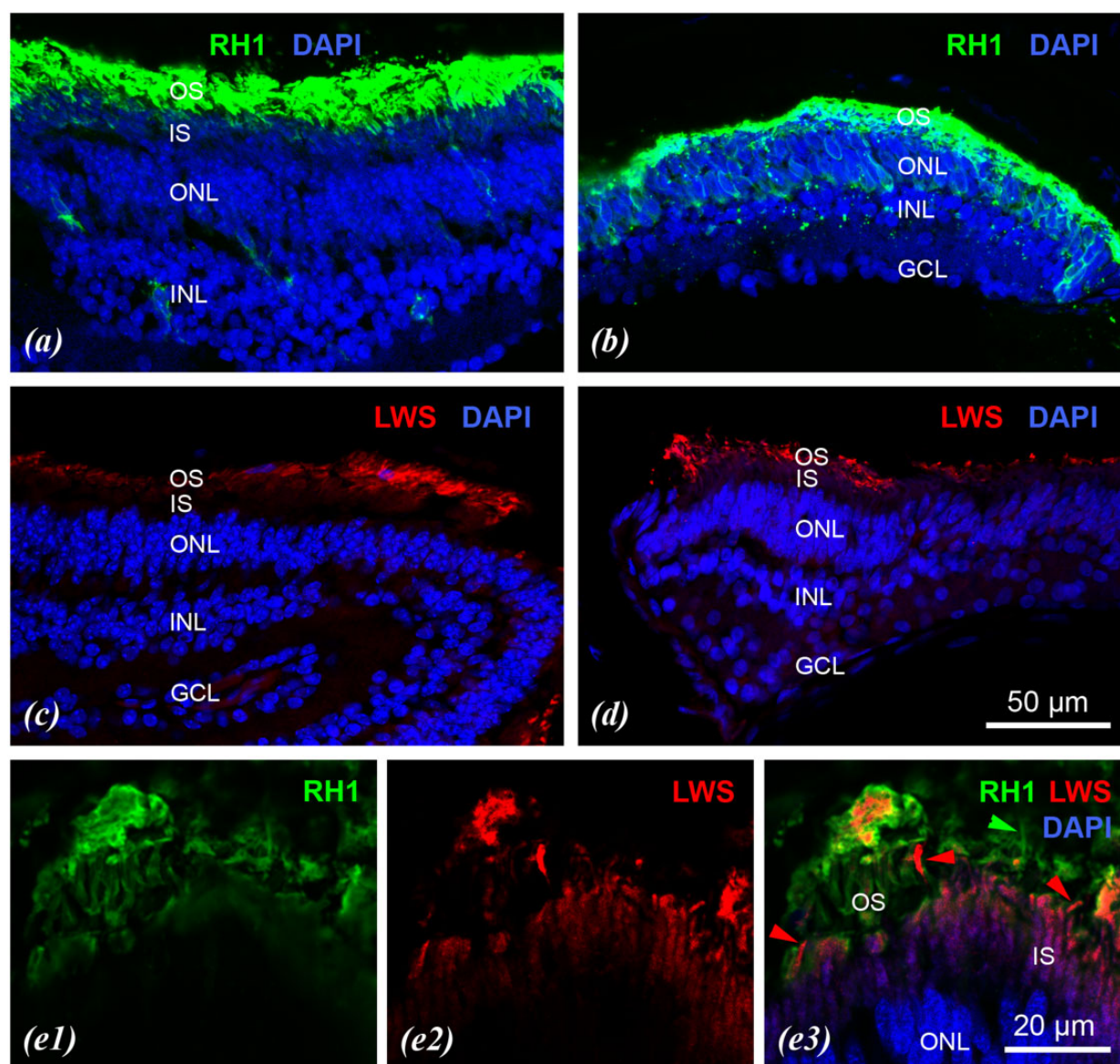


FIG. 2.—Retinal photoreceptor labelling in *Aniliios bituberculatus* (a, c) and *A. bicolor* (b, d, e). (a, b) Immunolabeling for RH1 in green. The transverse sections show strong labeling of rod outer segments, indicating a high rod density. Some of the rod somata in the ONL show fainter labeling. (c, d) Immunolabeling for LWS in red, with sections showing strong labeling of a substantial population of outer segments. (e) Double labeling of a section for RH1 (e1) and LWS (e2), the merge (e3) shows colocalization of the two opsins in many photoreceptor outer segments (yellow to orange), but also clear examples of outer segments showing exclusive RH1 label (green arrowhead) or LWS label (red arrowheads). The sections are counterstained with DAPI (blue) to show the retinal layers. All sections are oblique with artificially thicker nuclear layers, (c) shows a retinal fold. Overall, outer segment preservation is poor, many outer segments are ruffled, lumped, or torn. (a–d) Maximum-intensity projections of confocal image stacks; (e) single-focus image. OS, IS, photoreceptor outer and inner segments; ONL, outer nuclear layer; INL, inner nuclear layer; GCL, ganglion cell layer. Scale bar in (d) applies to (a–d), scale bar in (e3) applies to (e1–e3).

online). Furthermore, the PNA labeling was restricted to a band of punctate structures at the boundary between the ONL and photoreceptor inner segments, that is at the position of the external limiting membrane (supplementary fig. S5b, Supplementary Material online), it did not show the typical cone outer-segment labeling seen in other snakes.

Discussion

Our data and analytical results can be robustly interpreted as providing clear evidence for very substantial reduction of the photopic elements of the visual system of *Aniliios* spp. relative to that of the ancestral snake. Given that cone photoreceptors have not been reported for scolecophidians (e.g., Underwood

1967, 1970, 1977 and studies cited therein), that previous PCR surveys of retinal cDNA (Simões et al. 2015) failed to amplify any visual-opsin genes except *rh1* (typically expressed in rods and used in scotopic vision), and that MSP studies have found only a single, RH1-like pigment (Simões et al. 2015), we predicted that among vision genes that had been lost or become nonfunctional in scolecophidian genomes these would disproportionately have been linked functionally with photopic (cone-associated) rather than scotopic (typically rod-associated) vision, and this matches our results.

Our genome-wide sequencing data for *Anilius bituberculatus* are low coverage, but the fact that *Irat* (the shortest of the studied genes) is the only targeted vision gene that we could not find in the *A. bituberculatus* genome-wide sequencing data but that was present in the *A. bicolor* transcriptome encourages the interpretation that these genome-wide data are nonetheless informative for the purposes of this study. This is supported by lack of evidence for sequencing and/or assembly error. There is no evidence that the two vision genes absent in the *A. bicolor* transcriptome but present in the *A. bituberculatus* genome-wide data are pseudogenes in the latter—this pattern might also be explained by gene silencing in the eye and/or because of differences between the two species. Although we cannot be certain that the cone-phototransduction associated genes that we failed to find in the *A. bicolor* eye-transcriptome and the *A. bituberculatus* genome-wide data are all completely absent or nonfunctional in the genomes of these species (because we did not find the pseudogene remains of these genes), we are very confident that at least most are genuinely nonfunctional or absent, and we emphasize that none of the candidate losses (relative to the inferred ancestral snake) occurs among exclusively rod-associated phototransduction genes. Although we cannot rule out that any vision genes absent in the *Anilius* genome-wide and eye transcriptome data are pseudogenes, we predict that few if any of them are. This is because we have no evidence for any pseudogenization, and because the loss of these genes from the genome is plausible, given that extant scolecophidians all have a small-eyed, highly fossorial ecomorphology and they likely diverged from closest extant relatives (Alethinophidia) that retain these genes in the Cretaceous, more than 65 Ma (e.g., Burbrink et al. 2020). It is noteworthy that the vast majority of the genes potentially lost in *Anilius* spp. are also absent in highly fossorial mammals with reduced visual systems (fig. 1; Emerling and Springer 2014), suggestive of adaptive convergence. Thus, we can be confident that: 1) multiple visual genes likely present in the ancestral snake have been lost in *Anilius*, and 2) that loss is almost exclusively associated with genes mediating photopic vision. The *Anilius* visual system is clearly adapted for low-light conditions. These conclusions are robust despite the limitations that, beyond some previously reported *rh1* sequences (Simões et al. 2015), data on vision-related genes in scolecophidians are restricted currently to the eye transcriptome of *A. bicolor*

(based on ~80M raw reads) and the low-coverage genome-wide data for *A. bituberculatus* reported here.

The presence of *lws* in the *Anilius bituberculatus* genome-wide data and *A. bicolor* eye transcriptome (and of its product LWS in immunohistochemical preparations of their retinas) raises the question as to whether *lws* is present also in the genomes of the scolecophidians *Liotyphlops beui*, *Typhlophis squamosus*, *Amerotyphlops brongersmianus*, and *Habrophallos collaris*, with a previous failure to detect this gene in retinal cDNA (Simões et al. 2015) perhaps explained by its non- or very low-level expression in the eye. It might be noted that *lws* is generally the most persistent of the ancestral vertebrate photopic visual-opsin genes in cases where some of these genes are lost (e.g., Davies et al. 2012). Despite the expression of *lws* in photoreceptors, our immunohistological data agree with previous microscopic studies (Underwood 1967, 1970, 1977) in finding no evidence for morphologically cone-like photoreceptors in scolecophidians. A functional *lws* in taxa living primarily in dark environments might have a nonvisual function (Nei et al. 1997), perhaps in the control of circadian rhythms, and the expression of this gene in the pineal gland as well as the eye of at least the lizard *Anolis carolinensis* (Kawamura and Yokoyama, 1997) is consistent with this. If *lws* in *Anilius* does retain a visual function, it is perhaps integrated in the rod pathway given that many of the cone-pathway genes have been lost. Expression of rod visual-opsin genes in cone-like photoreceptors and *vice versa* has been reported in some vertebrates (e.g., Kojima et al., 2017; Simões et al. 2016b; Schott et al., 2016; Bhattacharyya et al. 2017) In this context, the identity and role of the photoreceptors expressing *lws* need further analysis.

Concerning the majority of *Anilius* photoreceptors that coexpress *rh1* and *lws*, we are not aware of the coexpression of *rh1* and typical cone opsin genes in single photoreceptors of any other vertebrates (e.g., Simões et al. 2016b; Schott et al. 2016). Coexpression of two or even three cone opsins in particular cones has been found in various vertebrate species from fish to mammals (e.g., Lukáts et al. 2005; Dalton et al. 2014; Hunt and Peichl, 2014; Isayama et al. 2014), but the coexpression of a rod and a cone opsin appears unique.

Despite the unexpected detection of functional *lws* in *Anilius* spp., the molecular data for this genus are consistent with anatomical data (e.g., Underwood 1977) that demonstrate that the scolecophidian visual system is simplified (reduced) in comparison to that of other (alethinophidian) snakes. Alethinophidians themselves have a simplified version of the genetic composition of the visual system that was likely present in the ancestral squamate. Figure 3 shows our interpretation of visual-gene loss among the squamates sampled in this study. Thus far, scolecophidians (at least *Anilius* spp.) are the only lepidosaurian reptiles presenting evidence for loss of the cone phototransduction genes *arr3*, *gnat2*, *gnpt2*, *pde6c*, and *gngb3*, the cone and rod phototransduction

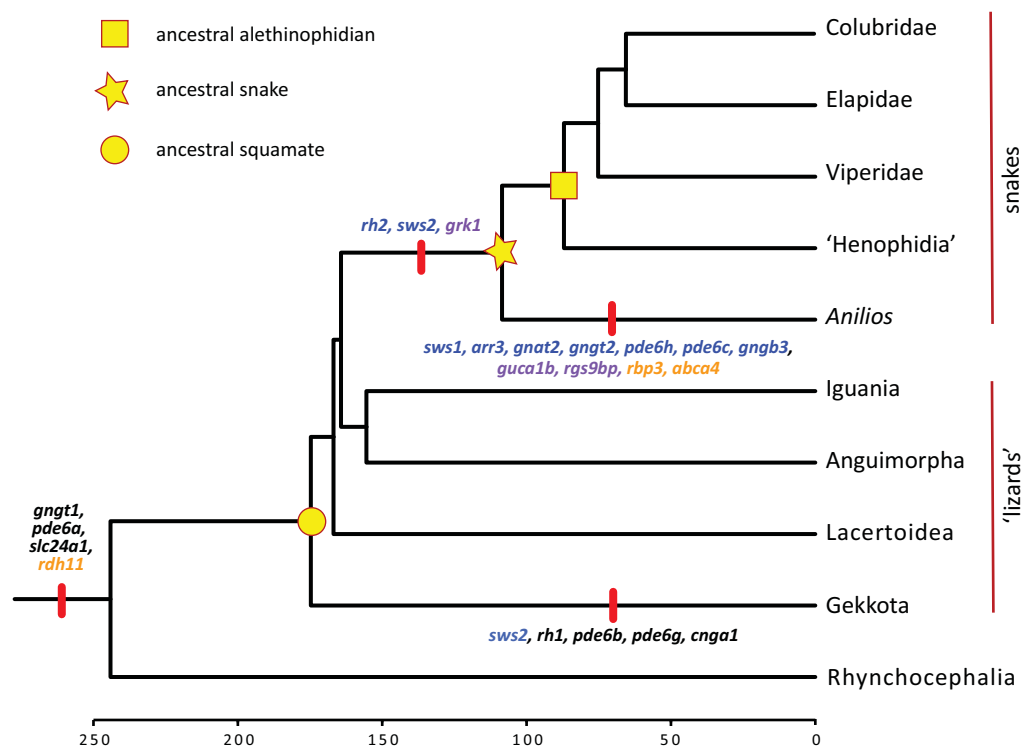


Fig. 3.—Proposed pattern of visual-gene loss during lepidosaurian evolution. Losses are indicated by red transverse bars on phylogenetic branches for cone (blue), rod (black), and cone and rod (purple) phototransduction genes, and visual cycle (orange) genes. Lower scale indicates approximate inferred age of phylogenetic divergences in millions of years. Phylogenetic relationships and mean estimates of divergence age estimates from www.vertlife.org.

genes *guca1b* and *rgs9bp*, and any visual-cycle genes other than *rdh11* (fig. 3; [supplementary tables S4 and S5, Supplementary Material online](#)). Figure 3 summarizes three main episodes of vision gene loss within Squamata; the loss of mostly rod phototransduction genes in presumably highly photopic stem geckos, the loss of three cone phototransduction genes in presumably scotopic stem snakes, and the very substantial loss of photopic transduction and visual-cycle genes in Scolecophidia (at least *Aniliios*).

Recently, [Miralles et al. \(2018\)](#) provided strong evidence for scolecophidian paraphyly, and they argued on this basis that a “scolecoid” phenotype, including reduced eyes without cone photoreceptors, was likely characteristic of the ancestral snake. In contrast, [Simões et al. \(2015\)](#) argued that absence of two visual-opsin genes (*sws1*, *lws*) in addition to the morphological simplifications in scolecophidian eyes made them a poor model for the nature of the ancestral snake eye, whether or not Scolecophidia is paraphyletic. Although we now know that *lws* is not absent in (all) scolecophidians, the evidence available thus far (presented here) adds substantial weight to the argument presented by [Simões et al. \(2015\)](#), in that it is highly unlikely that the ancestral snake lost functional copies of a large number of vision (especially cone phototransduction) genes that were subsequently re-evolved within Alethinophidia. Exceptions to Dollo’s law of evolutionary

irreversibility for loss of functional genes are considered extremely unlikely (e.g., [Collin and Miglietta 2008](#)), especially over longer timescales and for instances involving multiple genes. Available data indicate that the morphology and molecular machinery of the visual system of extant scolecophidians are substantially apomorphically reduced from that likely present in the ancestral snake. Clearly, examining the molecular genetic and anatomical components of a wider sample of scolecophidians would help to further advance this debate.

Materials and Methods

Genome-Wide Data

The new genome-wide data analyzed here are for a single individual of the Australian typhlopoid scolecophidian *Aniliios bituberculatus* (Peters 1863), obtained from the pet trade in 2009. There is no voucher specimen, but DNA sequences of mitochondrial and nuclear systematic markers very closely match those available on GenBank for *A. bituberculatus* ([supplementary appendix S1, Supplementary Material online](#)). All animal procedures complied with University of Leiden ethical guidelines. Genomic DNA was isolated from blood using a Qiagen Blood and Tissue DNeasy kit. A descriptive summary of the sequencing is reported in [supplementary table S1](#),

Supplementary Material online. A paired-end sequencing library with a target insert size of 500 bp was prepared from 5 µg DNA using a Paired-End Sequencing Sample Prep kit (Illumina Inc., San Diego). After amplification the library was analyzed on a Bioanalyzer 2100 DNA 1000 series II chip (Agilent, Santa Clara). The library was sequenced on an Illumina GAIIx instrument (3 lanes, 151 nucleotides paired-end), yielding a total of 26.2 Gbp of sequencing data. The reads were preprocessed to remove low-quality basecalls and adapter contamination, and assembled using the CLC Assembly Cell version 3.2 (CLC Bio, Aarhus) with a *k*-mer size of 25. This yielded a 1.69 Gbp draft genome assembly with a contig N50 of 1,917 nucleotides. Aligning the original sequencing data to the assembly shows that coverage is 7x. For comparative analyses, we downloaded DNA sequence data from NCBI (ncbi.nlm.nih.gov) for nine publicly available snake genomes, all of them alethinophidians (nonscolecophidians), nine lizards, and the tuatara *Sphenodon punctatus* ([supplementary table S2, Supplementary Material online](#)).

Eye-Transcriptomic Data

RNA later-preserved eyes of one specimen of *Anilius bicolor* ([supplementary appendix S2, Supplementary Material online](#)) were macerated in TRIzol and RNA purified using the PureLink™ RNA Mini Kit (Life Technologies/Ambion) using the manufacturer's protocol. The mRNA-Seq library was double-indexed and prepared with the mRNA-Seq Library Prep Kit v2 (Lexogen) following the manufacturer's protocol. It was sequenced with 56 other libraries in equimolar concentrations in one lane of an Illumina NovaSeq. A descriptive summary of the sequencing is reported in [supplementary table S1, Supplementary Material online](#). Low-quality reads were identified using Trimmomatic (Bolger et al. 2014) and removed. The remaining 59,727,054 read pairs were assembled using Trinity (Haas et al. 2013). Transcripts assembled by Trinity were then examined in Transdecoder (github.com/TransDecoder/TransDecoder/wiki) to predict ORFs and long coding regions. Transcripts were BLASTed (e-value of 1e-5) (Camacho et al. 2009) against the genome of *Python bivittatus* (GCA_000186305.2) and the Uniprot database (The UniProt Consortium 2018), and then annotated with Trinotate (Bryant et al. 2017). For comparative analyses, we downloaded data from publicly available eye transcriptomes of six snakes and of six lizards ([supplementary table S2, Supplementary Material online](#)).

Molecular Data Analyses

We selected 48 genes of interest for visual-system biology, focusing on the retinal phototransduction and visual-cycle (cascade) genes that were also analyzed by Emerling and Springer's (2014) study of regression of the visual system in fossorial mammals. Each of the genomes was interrogated by the Basic Local Alignment Search Tool, BLAST (Camacho et al. 2009) under

tblastn, once for each of the 48 vision genes, using small (two or three-taxon) seed sequence sets. Sequences with low BLAST e values (<e-20) were then subjected to a second BLAST analysis using tblastn against vertebrate entries in NCBI's non-redundant protein database to verify that they were not homologous with other genes not under consideration. Following this, the BLAST e-value threshold was lowered (<e-7) and newly identified sequences were subjected to the same retesting process to greatly reduce the chance that highly divergent sequences were missed due to low similarity to the limited seed data set. These results were then supplemented by NCBI searches of the candidate genes to ascertain whether sequences not recovered during the BLAST searches had nonetheless been generated for individuals of the same species.

For sequences recovered by BLAST that belong to gene families of multiple genes having high similarity (GNGT*, GNAT*, CNGA*, CNGB*, GRK*, GUCA* GUCY*, RDH*, RBP*) phylogenetic analyses were used to further assess the identification of the 48 vision genes. Vision genes that had sequences recovered by BLAST (and verified by phylogenetic analysis) that lacked unexpected stop codons and/or frameshift indels in the consensus regions of the protein (as determined from the Prosite and EMBL-EBI databases: prosite.expasy.com and ebi.ac.uk, respectively) were interpreted as present and potentially functional.

Visual-opsin gene sequences identified in the *Anilius bicolor* transcriptome and *A. bituberculatus* genome-wide data were added to a database comprising publicly available snake and lizard visual-opsin gene sequences, and were aligned using default settings in MAFFT (Katoh et al. 2002). Alignments were inspected by eye for nucleotides and amino acids, and adjusted manually to ensure nucleotides were in-frame and that indels did not include partial codons. Trees were inferred with RAxMLv8 (Stamatakis 2014) using majority rule bootstrapping (Pattengale et al. 2010), randomized MP starting trees, and a fast hill-climbing algorithm, under the GTR+G+I model (determined as best-fitting for all visual opsin alignments using jModelTest 2: Durrbin et al. 2012). Visual-opsin gene trees were rooted using *lws* based on previous studies of opsin gene phylogenetics (e.g., Pisani et al. 2006; Feuda et al. 2012), with individual opsin trees rooted with *Sphenodon*.

Ancestral-state estimation was undertaken using parsimony, with putative gene losses considered irreversible. Functional vision genes were considered to be present in the ancestral snake and the ancestral alethinophidian (i.e., the ancestor of all nonscolecophidian snakes) where they are present in the genomes of at least one alethinophidian and one lizard, with absences in *Anilius* interpreted as possible loss in at least this scolecophidian lineage.

Retinal Immunohistochemistry

Whole heads of two *Anilius bituberculatus* and one *A. bicolor* ([supplementary appendix S3, Supplementary Material online](#)) were fixed in 4% (w/v) paraformaldehyde in 0.01 M

phosphate-buffered saline (PBS, pH 7.4) for less than 1 day, then stored in PBS with 0.05% sodium azide at 4 °C. Later, eyes were removed and stored in the same buffer. The very small eyes were measured with calipers, but orientation could not be recorded. Opsin immunohistochemistry was performed on frozen transverse sections of eyes. Whole eyes were cryoprotected by successive immersion in 10%, 20%, and 30% (w/v) sucrose in 0.1 M phosphate buffer (PB, pH 7.4) and transferred to Tissue Freezing Medium (Leica Biosystems, Wetzlar, Germany). Eyes were embedded and frozen in blocks, oriented to obtain sections parallel to the anterior–posterior axis, sectioned to 16 μm with a cryostat, and collected on SuperFrost slides (Menzel GmbH, Braunschweig, Germany). Due to the small eye size, only the most central sections yielded good transverse aspects of the retina, in more peripheral sections the retina was sectioned more obliquely.

Immunolabeling followed protocols previously used to label opsins in snake retinae (Hauzman et al. 2014; Gower et al. 2019). Rod opsin RH1 was detected with the mouse monoclonal antibody rho4D2 (dilution 1:1,000; kindly provided by R. S. Molday; Hicks and Molday 1986). Cone opsins were detected with the rabbit antiserum JH492 against the longwave-sensitive LWS cone opsin (dilution 1:2,000; kindly provided by J. Nathans; Wang et al. 1992), and with the goat antiserum sc-14363 against the shortwave-sensitive SWS1 cone opsin (dilution 1:500; Santa Cruz Biotechnology Inc., Heidelberg, Germany). Binding sites of primary antibodies were visualized by indirect immunofluorescence. Omission of primary antibodies from incubation solutions resulted in no staining. Double-labeling was done by incubating sections in a mixture of primary and secondary antibodies. For some sections, staining with Alexa 647-conjugated peanut agglutinin (PNA; dilution 1:250; Invitrogen) was added to the opsin immunolabeling. PNA is a cone marker in various vertebrates, including snakes (Hauzman et al. 2017; Gower et al. 2019). Sections were counterstained with 4,6-diamidino-2-phenylindole (DAPI) to visualize retinal nuclear layers. One unsectioned *A. bituberculatus* retina was immunolabeled for cone opsins free-floating, then flat-mounted on a slide with the photoreceptor side up. Here incubation was in a mixture of the cone opsin antisera JH492 and sc-14363 for three days at room temperature, incubation in the secondary antiserum mixture was for 1 h.

Stained tissue was analyzed with a laser scanning microscope (LSM) Olympus FluoView 1000 using the FV 1.7 software (Olympus). LSM images and z-stack projections were examined with ImageJ (<http://imagej.nih.gov/ij/>). Images for illustration were adjusted for brightness and contrast using Adobe Photoshop.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

DJG thanks Mark Wilkinson for help with probability calculation, and Jeff Streicher for critiquing an earlier draft. BFS thanks Mark Hutchinson (South Australia Museum) for help with fieldwork and Alastair Ludington for bioinformatic support. The authors thank two anonymous reviewers for constructively critical reviews of the submitted manuscript. JFF's contribution was supported by a NERC GW4+ PhD studentship and by Japan Society for the Promotion of Science Grant-in-Aid KAKENHI grant 18F18788. BFS was supported by the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie fellowship 703438, and BFS, DP, and KLS are funded by an Australian Research Council Discovery Grant DP180101688. HMIK was supported by the Top Sector Syngenepep grant (731.014.206) from the Netherlands Organization for Scientific Research (NOW), and by Naturalis Biodiversity Center. The research was also supported by Leverhulme Trust grant RPG-342 (to DJG, Nathan S. Hart, Julian C. Partridge, and David M. Hunt).

Authors' Contributions

JFF and BFS carried out the vision gene bioinformatics (helped by CVH) and phylogenetics. FJV, HMIK, NRC, CVH, and MKR contributed genome-wide sequencing; BFS and KLS eye-transcriptome sequencing. LP and SM generated and analyzed immunohistochemistry data. DJG conceived of and designed the study with help from JFF, DP, MKR, and BFS. All authors contributed to writing.

Data Availability

The data underlying this article are available in the article, in its online [supplementary material](#), and GenBank accessions MK895004–031.

Literature Cited

- Bellairs ADA. 1972. Comments on the evolution and affinities of snakes. In: Joysey KA, Kemp TS, editors. *Studies in vertebrate evolution*. Edinburgh: Oliver & Boyd. p. 157–172.
- Bellairs ADA, Underwood G. 1951. The origin of snakes. *Biol Rev Camb Philos Soc.* 26(2):193–237.
- Bhattacharyya N, Darren B, Schott RK, Tropepe V, Chang BSW. 2017. Cone-like rhodopsin expressed in the all-cone retina of the colubrid pine snake as a potential adaptation to diurnality. *J Exp Biol.* 220(Pt 13):2418–2425.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina Sequence Data. *Bioinformatics* 30(15):2114–2120.
- Bryant DM, et al. 2017. A tissue-mapped Axolotl de novo transcriptome enables identification of limb regeneration factors. *Cell Rep.* 18(3):762–776.
- Burbrink FT, et al. 2020. Interrogating genomic-scale data for Squamata (lizards, snakes and amphisbaenians) shows no support for key traditional morphological relationships. *Syst Biol.* 69(3):502–520.
- Caldwell MW. 2007. Snake phylogeny, origins, and evolution: the role, impact, and importance of fossils 1869–2006. In: Anderson JS, Sues H-

- S, editors. Major transitions in vertebrate evolution. Bloomington and Indianapolis (IN): Indiana University Press. p. 253–302.
- Camacho C, et al. 2009. BLAST+: architecture and applications. *BMC Bioinformatics*. 10:421.
- Caprette CL, Lee MSY, Shine R, Mokany A, Downhower JF. 2004. The origin of snakes *Serpentes* as seen through eye anatomy. *Biol J Linn Soc*. 81(4):469–482.
- Castoe TA, et al. 2013. The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. *Proc Natl Acad Sci U S A*. 110(51):20645–20650.
- Collin R, Miglietta MP. 2008. Reversing opinions on Dollo's Law. *Trends Ecol Evol*. 23(11):602–609.
- Da Silva FO, et al. 2018. The ecological origins of snakes as revealed by skull evolution. *Nat Commun*. 9(1):376.
- Dalton BE, Loew ER, Cronin TW, Carleton KL. 2014. Spectral tuning by opsin coexpression in retinal regions that view different parts of the visual field. *Proc R Soc B*. 281(1797):20141980.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods*. 9(8):772.
- Darwin C. 1859. On the origin of species by means of natural selection. London: J. Murray.
- Davies WL, et al. 2009. Shedding light on serpent sight: the visual pigments of henophidian snakes. *J Neurosci*. 29(23):7519–7525.
- Davies WL, Collin SP, Hunt DM. 2012. Molecular ecology and adaptation of visual photopigments in craniates. *Mol Ecol*. 21(13):3121–3158.
- Eme L, Spang A, Lombard J, Stairs CW, Ettema TJ. 2017. Archaea and the origin of eukaryotes. *Nat Rev Microbiol*. 15(12):711–723.
- Emerling CA. 2017. Genomic regression of claw keratin, taste receptor and light-associated genes provides insights into biology and evolutionary origins of snakes. *Mol Phylogenet Evol*. 115:40–49.
- Emerling CA, Springer MS. 2014. Eyes underground: regression of visual protein networks in subterranean mammals. *Mol Phylogenet Evol*. 78:260–270.
- Feuda R, Hamilton SC, McInerney JO, Pisani D. 2012. Metazoan opsin evolution reveals a simple route to animal vision. *Proc Natl Acad Sci U S A*. 109(46):18868–18872.
- Gower DJ, et al. 2019. Evolution of the eyes of vipers with and without infrared-sensing pit organs. *Biol J Linn Soc*. 126(4):796–823.
- Haas BJ, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc*. 8(8):1494–1512.
- Harrington SM, Reeder TW. 2017. Phylogenetic inference and divergence dating of snakes using molecules, morphology and fossils: new insights into convergent evolution of feeding morphology and limb reduction. *Biol J Linn Soc*. 121(2):379–394.
- Hauzman E, et al. 2014. Comparative study of photoreceptor and retinal ganglion cell topography and spatial resolving power in Dipsadidae snakes. *Brain Behav Evol*. 84(3):197–213.
- Hauzman E, Bonci DMO, Suárez-Villota EY, Neitz M, Ventura DF. 2017. Daily activity patterns influence retinal morphology, signatures of selection, and spectral tuning of opsin genes in colubrid snakes. *BMC Evol Biol*. 17(1):249.
- Hicks D, Molday RS. 1986. Differential immunogold dextran labeling of bovine and frog rod and cone cells using monoclonal-antibodies against bovine rhodopsin. *Exp Eye Res*. 42(1):55–71.
- Holland LZ, et al. 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res*. 18(7):1100–1111.
- Hsiang AY, et al. 2015. The origin of snakes: revealing the ecology, behavior, and evolutionary history of early snakes using genomics, phenomics, and the fossil record. *BMC Evol Biol*. 15:87.
- Hunt DM, Peichl L. 2014. S cones: evolution, retinal distribution, development, and spectral sensitivity. *Vis Neurosci*. 31(2):115–138.
- Ishiyama T, et al. 2014. Coexpression of three opsins in cone photoreceptors of the salamander *Ambystoma tigrinum*. *J Comp Neurol*. 522(10):2249–2265.
- Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 30(14):3059–3066.
- Kawamura S, Yokoyama S. 1997. Expression of visual and nonvisual opsins in American chameleon. *Vision Res*. 37(14):1867–1871.
- Kemp TS. 2015. The origin of higher taxa: palaeobiological, developmental and ecological perspectives. Oxford: Oxford University Press.
- Kojima K, et al. 2017. Adaptation of cone pigments found in green rods for scotopic vision through single amino acid mutation. *Proc Natl Acad Sci U S A*. 114(21):5437–5442.
- Lane N. 2015. The vital question: energy, evolution, and the origins of complex life. London: WW Norton & Company.
- Lee MS. 2005. Molecular evidence and marine snake origins. *Biol Lett*. 1(2):227–230.
- Lee MSY, Hugall AF, Lawson R, Scanlon JD. 2007. Phylogeny of snakes (*Serpentes*): combining morphological and molecular data in likelihood, Bayesian, and parsimony analyses. *Syst Biodivers*. 5(4):371–389.
- Lukáts Á, Szabó A, Röhlich P, Vígh B, Szél Á. 2005. Photopigment coexpression in mammals: comparative and developmental aspects. *Histol Histopathol*. 20(2):551–574.
- Miralles A, et al. 2018. Molecular evidence for the paraphyly of Scolecophidia and its evolutionary implications. *J Evol Biol*. 31(12):1782–1793.
- Nei M, Zhang J, Yokoyama S. 1997. Color vision of ancestral organisms of higher primates. *Mol Biol Evol*. 14(6):611–618.
- Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. 2010. How many bootstrap replicates are necessary? *J Comput Biol*. 17(3):337–354.
- Perry BW, et al. 2018. Molecular adaptations for sensing and securing prey and insight into amniote genome diversity from the garter snake genome. *Genome Biol Evol*. 10(8):2110–2129.
- Pisani D, Mohun SM, Harris SR, McInerney JO, Wilkinson M. 2006. Molecular evidence for dim-light vision in the last common ancestor of the vertebrates. *Curr Biol*. 16(9):R318–R319.
- Rieppel O. 1988. A review of the origin of snakes. In: Hecht MK, Wallace B, Prance GT, editors. *Evolutionary biology*. Boston (MA): Springer. p. 37–130.
- Schott RK, et al. 2016. Evolutionary transformation of rod photoreceptors in the all-cone retina of a diurnal garter snake. *Proc Natl Acad Sci U S A*. 113(2):356–361.
- Schott RK, Van Nynatten A, Card DC, Castoe TA, Chang BS. 2018. Shifts in selective pressures on snake phototransduction genes associated with photoreceptor transmutation and dim-light ancestry. *Mol Biol Evol*. 35(6):1376–1389.
- Simões BF, et al. 2015. Visual system evolution and the nature of the ancestral snake. *J Evol Biol*. 28(7):1309–1320.
- Simões BF, et al. 2016a. Visual pigments, ocular filters and the evolution of snake vision. *Mol Biol Evol*. 33(10):2483–2495.
- Simões BF, et al. 2016b. Multiple rod–cone and cone–rod photoreceptor transmutations in snakes: evidence from visual opsin gene expression. *Proc R Soc B*. 283(1823):20152624.
- Smith JM, Szathmari E. 1997. The major transitions in evolution. Oxford: Oxford University Press.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9):1312–1313.
- The UniProt Consortium. 2018. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*. 47:D506–D515.
- Underwood G. 1967. A contribution to the classification of Snakes. London: British Museum Natural History Publications.

- Underwood G. 1970. The eye. In: Gans C, Parsons TS, editors. *Biology of the Reptilia: morphology B*. New York: Academic Press. p. 1–97.
- Underwood G. 1977. Simplification and degeneration in the course of evolution of squamate reptiles. *Colloq Intx Cent Natl Rech Sci*. 266:341–353.
- Wang YS, et al. 1992. A locus-control region adjacent to the human red and green visual pigment genes. *Neuron*. 9(3):429–440.
- Wiens JJ, et al. 2012. Resolving the phylogeny of lizards and snakes Squamata with extensive sampling of genes and species. *Biol Lett*. 8(6):1043–1046.
- Yi H, Norell MA. 2015. The burrowing origin of modern snakes. *Sci Adv*. 1(10):e1500743.

Associate editor: Takashi Makino