Simon M. Tierney, Steven J. B. Cooper, Kathleen M. Saint, Terry Bertozzi, Josephine Hyde, William F. Humphreys, and Andrew D. Austin

**Opsin transcripts of predatory diving beetles: a comparison of surface and subterranean photic niches**

Royal Society Open Science, 2015; 2:140386-1-140386-12

©2015 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

Source:
http://rsos.royalsocietypublishing.org/content/2/1/140386

As per Creative Commons License: http://creativecommons.org/licenses/by/4.0/
You are free to:
Share copy and redistribute the material in any medium or format
Adapt remix, transform, and build upon the material for any purpose, even commercially.
The licensor cannot revoke these freedoms as long as you follow the license terms.

**January 30, 2015**

http://hdl.handle.net/2440/88988
1. Summary

The regressive evolution of eyes has long intrigued biologists yet the genetic underpinnings remain opaque. A system of discrete aquifers in arid Australia provides a powerful comparative means to explore trait regression at the genomic level. Multiple surface ancestors from two tribes of diving beetles (Dytiscidae) repeatedly invaded these calcrete aquifers and convergently evolved eye-less phenotypes. We use this system to assess transcription of opsin photoreceptor genes among the transcriptomes of two surface and three subterranean dytiscid species and test whether these genes have evolved under neutral predictions. Transcripts for UV, long-wavelength and ciliary-type opsins were identified from the surface beetle transcriptomes. Two subterranean beetles showed parallel loss of all opsin transcription, as expected under ‘neutral’ regressive evolution. The third species Limbodessus palmulaoides retained transcription of a long-wavelength opsin (lwop) orthologue, albeit in an aphotic environment. Tests of selection on lwop indicated no significant differences between transcripts derived from surface and subterranean habitats, with strong evidence for purifying selection acting on L. palmulaoides lwop. Retention of sequence integrity and the lack of evidence for neutral evolution raise the question of whether we have identified a novel pleiotropic role for lwop, or an incipient phase of pseudogene development.
2. Introduction

The evolutionary degeneration or loss of traits, known as regressive evolution, has generated considerable discussion in Lamarckian, Darwinian and Neutralist contexts, with eye-loss among subterranean animals being central to the debate (reviewed by [1–8]). Neutral evolutionary theory (non-adaptive evolution) predicts that if genes specifically responsible for vision are not under directional selection, then they should accrue random mutations through genetic drift, ultimately developing into pseudogenes given sufficient time [9–11]. There are very few empirical examples of such genetic loss of function in ‘eye genes’ [12], and instances where visual transduction genes continue to be transcribed in a lightless environment raise the possibility of pleiotropy [13,14]. However, the evolutionary origins of these study systems are rarely of sufficient palaeo-age to allow pseudogenes to develop or exhibit the diversity required for robust comparative phylogenetic assessment.

Tracking the history of an organism as it enters into and diversifies within a novel niche permits the investigation of a ‘natural experiment’ set up by evolution. Such studies are most powerful when contrasting closely related lineages with divergent phenotypes and when there are repeated independent origins of the focal trait [15]. The stygofauna (aquatic obligate subterranean fauna) of arid zone Australia represent a diverse assemblage of subterranean metazoans that derive from surface-water ancestors, which have independently colonized calcare aquifers during periods of continental aridification in the Plio-Pleistocene [16–20] and so the niche shift may be climate driven. One of the best examples of parallel evolution within this system is that of the predatory diving beetles (Dytiscidae) of the Yilgarn Craton in Western Australia (WA), represented by approximately 100 known stygobitic species from 45 isolated calcare aquifers [21–23]. A consistent trait among these stygobic beetles is the complete lack of eyes, lack of pigmentation and wing reduction ([24–26] and references therein). Phylogeographic studies have provided strong evidence that the calcare aquifers represent closed island systems, with no evidence for gene flow among stygobitic species from different aquifers and an absence of gene flow from surface species over millions of years [19–23,27], potentially long enough for neutral genetic changes to be detected in eye genes that are no longer functional [11]. These attributes, and the independent evolution of the majority of stygobitic beetle species from surface ancestors, provide a powerful model system to explore the long-term genomic changes that accompany regressive evolution of subterranean animals.

The transduction of photons into neural impulses enables animal vision. Visual pigments of photoreceptor cells absorb photons of specific wavelength sensitivities that activate phototransduction cascades. All visual pigments consist of an opsin apo-protein and a chromophore, and the point at which these molecules bind (Schiff-base linkage) determines peak wavelength sensitivity. Point mutations and gene duplication in opsin can tune the pigment to alternative wavelength sensitivity, as can rhabdomic filters. The shared ancestry of animal vision [28,29] has made opsin an ideal candidate gene for phylogenetic studies generally, as well as those that track species entry into novel photic niches (e.g. [30,31]). Non-visual opsins are also purported to play regulatory functions in light-mediated processes (e.g. circadian rhythm [32]), and detection of visual opsin expression in arthropod brains may indicate pleiotropic roles (reviewed by [33]).

The aim of this study is to explore the regressive evolution of opsin genes from a molecular standpoint. We take a qualitative candidate gene approach and use high-throughput sequencing techniques to compare transcriptomes of two surface and three subterranean dytiscid beetle species that are blind and lack pigmentation. We target opsin phototransduction genes and explore the relative rates of non-synonymous (dN) versus synonymous (dS) mutations in order to make inferences regarding evolutionary mode. Neutral theory predicts that photoreceptor genes, no longer required in stygobitic taxa, should evolve into pseudogenes. Hence, we should expect an absence of transcription of the ancestral gene (that of surface species) or non-directional rates of mutation (dN = dS) in genes that continue to be transcribed in an aphotic niche.

3. Material and methods

3.1. Taxon sampling

Diving water beetles from two tribes of Dytiscidae (Bidessini, Hydroporini) were assessed. Surface species include one bidessine, *Allodessus bistrigatus* (Clark, 1862), and one hydroporine, *Paroster nigroadumbratus* (Clark, 1862); both were collected in September 2012, near Forreston, South Australia (−34.6889°, 138.8933°). Subterranean beetles were collected from bore holes of the Yilgarn region,
WA, from which we sourced two bidessine species: *Limbodessus palmulaoides* (Watts & Humphreys, 2006), collected in August 2011, Laverton Downs calcrete, WA (−28.3983°, 122.2038°) and *Neobidessodes guttidgei* (Watts & Humphreys, 2003), collected in May 2012, Three Rivers Station calcrete, WA (−25.2786°, 119.1834°). Finally, one subterranean hydroporine, *Paroster macrosturtensis* (Watts & Humphreys, 2006), was collected in May 2012, from Sturt Meadows calcrete, WA (−28.7155°, 120.8931°). Specimens were collected under licences SF007802 & SF008440 (W. F. Humphreys), issued by the Department of Environment and Conservation. Each of these species have been the subject of phylogeographic studies [22,34], with further molecular ecological studies for two species (*L. palmulaoides* and *P. macrosturtensis*) [35–37] that have confirmed their long-term independent evolution from sympatric stygobitic species and surface relatives over millions of years. All specimens were preserved in RNAlater.

### 3.2. cDNA synthesis and high-throughput sequencing

Depending on taxon body size [25], 5–10 heads from each species were pooled and homogenized using steel bead disruption. Pooling individuals within a species effectively increases the power of results, particularly when a qualitative presence or absence of expression is desired. Total RNA was purified following standard protocols (RNeasy Plus Micro Kit; Qiagen). Double-stranded cDNA was synthesized and PCR-amplified using the SMARTer cDNA Synthesis Kit and Advantage 2 PCR Kit (Clontech). PCR optimization procedures were gel-verified. cDNA was sequenced on an Illumina HiSeq2000, generating either 100 or 150 bp paired-end reads with TruSeq adapters.

### 3.3. Bioinformatics

#### 3.3.1. Quality control

Raw reads were assessed using FASTQC (Babraham Institute) and then edited and trimmed in CUTADAPT [38] to filter-out: low-quality ends of reads (phred scores <30), TruSeq barcoded adapters, SMARTer adapters, poly-A tails and sequences less than 25 bp.

#### 3.3.2. Assembly

Transcripts were de novo assembled using the Trinity platform’s three-step process, which assembles then aligns contigs via probability derived de Bruijn graphs [39]. Assembled contigs were quality assessed with BOWTIE, aligning them back to raw reads to indicate the proportion of proper-paired reads obtained. Read coverage depth for transcripts of interest were ascertained with BEDTOOLS (v. 2.22.0) and summarized in SPSS (v. 21).

#### 3.3.3. Orthologue search

We used BLASTx to conduct amino acid translated queries of Trinity derived transcripts against a Candidate Set of 16 target proteins (listed in figure 1), for non-visual opsins (ciliary-type and pteropsin) and visual opsins (UV, blue and long-wavelength sensitive), derived from: sunburst diving beetle *Thermonectus marmoratus* (Coleoptera: Dytiscidae); red flour beetle *Tribolium castaneum* (Coleoptera: Tenebrionidae); European honeybee *Apis mellifera* (Hymenoptera: Apidae) and Chinese swallowtail butterfly *Papilio xuthus* (Lepidoptera: Papilionidae). Target sequences were obtained from GenBank, accessed June 2013 (figure 1 and electronic supplementary material, tables S1 and S2). We then performed a reciprocal search (tBLASTn), followed by BLASTn searches on the best contig hit (derived from BLASTx) for each candidate protein. An orthologous match was considered positive when: alignment length was greater than or equal to 50%; protein identities were greater than or equal to 30% and nucleotide identities were greater than or equal to 70% [40]. Failing this stringency, BLAST2BLASTn alignments for sequence-pairs were undertaken between best-hit contigs from our study that positively matched an opsin orthologue (subject sequence) and best-hit contigs that failed to match any opsin orthologues on GenBank (query sequence). Best-hit dytiscid transcripts that resulted in a positive orthologue match were deposited in GenBank (accession numbers KP219380–KP219386). A conceptual workflow of these methods is outlined in the electronic supplementary material, figure S1.
3.3.4. Phylogenetics

The Candidate Set opsin amino acid sequences (excluding unalignable putative UTR) were first aligned with CLUSTAL OMEGA [41]; best-hit dytiscid transcripts were manually aligned to the nucleotide translation in SE-AL v. 2.0a11 and reading frames were double checked in MACCLADE 4.06. MRBAYES v. 3.2 was used to construct a 50% majority rule tree, derived from two independent runs of four chains (one cold). We used an objective procedure (least constrained model with ample generation time [42]), to avoid potential errors related to incorrect a priori model selection. An A GTR + I + Γ model was applied to each codon partition, with non-informative priors, unlinked parameters and variable rates permitted. Run length was determined by: (i) likelihood plots, (ii) the average standard deviation of split frequencies, and (iii) estimated sample size. We used a relative burn-in to discard the initial 25% of samples.

3.3.5. Tests of selection

We calculated P-distances for amino acid and nucleotide sequences (± s.e). Pairwise Z-tests (Nei–Gojobori method) of neutral evolution, purifying selection and positive selection were undertaken in MEGA 5.2.2, with pairwise deletion of ambiguous sites and 500 bootstrap replicates to estimate variance. We then used HYPhy 2.1 to undertake phylogenetic hypothesis testing using Bayesian inferred trees. We used likelihood ratio tests (LRT) to explore: (i) pairwise relative rates of evolution, (ii) tree-wide global rates of selection versus local rates, (iii) comparison of rates among select branches of the tree, as well as (iv) two site-specific LRT methods that calculate dN/dS independently at each codon, which differ in analytical robustness: single-likelihood ancestor counts (SLAC) are simplistic; fixed effect likelihood (FEL) methods are more detailed and less susceptible to Type I errors compared with comparable procedures [43].

4. Results

4.1. Transcriptome assembly

Metrics for the de novo assemblies are summarized in the electronic supplementary material, table S3. We recovered between 50.6 and 78.9 million paired-end reads for each of the five species. Assembled
transcripts ranged from 84,667 to 182,240 per species (contig N50 ranged from 565 to 1067 bp) and total aligned reads ranged from 37 to 74.3 million reads.

4.2. Searches for orthologous opsin sequences in the dytiscid beetle transcripts

The complete BLAST results are provided in the electronic supplementary material, table S1 and a qualitative representation (presence/absence of orthologous transcript) comparing species is presented in figure 1. Overall, positive matches to 14 of the 16 Candidate opsin genes were recovered. Blue-sensitive target candidates were the only opsin class not detected at all. Transcripts from both surface beetle species resulted in positive matches to the two non-visual opsin candidates, pteropsin and c-opsin, and all five ultraviolet opsin genes (uvop). Each of the seven long-wavelength opsin (lwop) candidates registered matches in both surface species and also in one subterranean species (L. palmulaoides). Transcripts from the remaining two subterranean species, N. gutteridgei and P. macrosturtensis, did not match any of the 16 Candidate opsin genes. Results from surface species are presented independently of subterranean species (below) and correspond to results tabulated in the electronic supplementary material, table S1.

4.2.1. Surface beetle orthologues

Non-visual opsin orthologues were identified among other non-Candidate insect taxa (electronic supplementary material, tables S1.1.1, S1.1.3, S1.4.1, S1.4.3). Of these, the best orthologous match was to a parapinopsin gene of the silkworm Bombyx mori (XM_004928326). For the visual opsins, there were amino acid and nucleotide matches to uvop of all beetle Candidates and the honeybee Candidate (electronic supplementary material, tables S1.1.1, S1.1.2, S1.4.1, S1.4.2). The best uvop nucleotide matches were to T. marmoratus, with A. bistrigatus best matched to T. marmoratus uvop-2 (electronic supplementary material, table S1.1.3), and P. nigroadumbratus best matched to T. marmoratus uvop-1 (electronic supplementary material, table S1.4.3). Both surface dytiscid beetle species returned amino acid matches to blue opsin Candidates; however, nucleotide searches of these best-hit transcripts did not match to any blue opsin in GenBank, and the best matches were to uvop of T. marmoratus. Long-wavelength opsin Candidates yielded the highest scores of all our BLAST orthologue searches. Both focal surface dytiscid beetle species transcripts returned the highest scoring nucleotide matches (BLASTn, electronic supplementary material, tables S1.1.2, S1.4.2), which were most closely affiliated to lwop of the North American dytiscid beetle T. marmoratus.

4.2.2. Subterranean beetle orthologues

A single transcript from the L. palmulaoides assembly rated as the best hit for all opsin Candidates in the amino acid search (electronic supplementary material, table S1.2.1) and the best nucleotide match to this transcript was to T. marmoratus lwop (electronic supplementary material, table S1.2.3). No positive matches to any of the opsin Candidates were found for the two subterranean species N. gutteridgei and P. macrosturtensis. However, for these two species, the majority of GenBank nucleotide matches (BLASTn derived) to best-hit transcripts (derived from BLASTx) were to other members of the G protein-coupled receptor family (which includes opsins) that were not related to phototransduction (electronic supplementary material, tables S1.3.3, S1.5.5.3). This suggests that our assemblies were of high quality and capable of discerning transcripts at high fidelity.

4.3. Orthologue coverage depth

Depth of sequencing coverage for each of the seven identified opsin orthologues (un-edited) is summarized in table 1 and plotted on a lognormal scale in figure 2. The non-visual c-opsin from surface species exhibited the lowest levels of mean coverage per nucleotide base (less than 12 times coverage). The visual opsins of surface beetle species showed coverage at orders of magnitude higher (uvop means less than 1220 times coverage; lwop means less than 69,015 times coverage). The lwop of the subterranean L. palmulaoides displayed a mean sequence coverage depth of 22.51 (±s.e. 0.57) per nucleotide base, which is double the level of coverage found for c-opsin in the two surface species; this suggests that our finding of the expression of a visual opsin in a blind subterranean beetle is not an aberration.

4.4. Phylogenetic assessment of predicted opsin transcripts

All best-hit transcripts that returned positive matches to opsins in the Candidate Set, and detailed in the BLAST results above, were then placed within a comparative phylogenetic framework. We
augmented the 16 candidate opsin matrix with an additional three non-visual opsins (figure 3; electronic supplementary material, table S2). Based on the aforementioned BLAST results, we included the parapinopsin of B. mori. To root the tree, we included two vertebrate non-visual opsins as outgroups: multiple tissue opsin of Takifugu rubripes and teleost multiple tissue opsin of Danio rerio. An aligned nucleotide matrix comprising the Candidate Set opsins and their dytiscid beetle best-hit positive match transcripts (26 sequences, 1005 nucleotides [44]) was subjected to a Bayesian phylogenetic analysis of 10 million iterations.

The resultant Bayesian tree (with posterior probability node support) corroborates BLAST search results and is presented in figure 3. All opsin classes (ciliary, UV, blue and long-wavelength) formed fully resolved monophyletic groups with maximal posterior probability support. The BLAST-predicted
Figure 3. Opsin phylogeny. Majority rule consensus tree of visual and non-visual opsin photoreceptors derived from Bayesian inference, with posterior probability node support. Tree branches and terminal taxa are colour coded to identify Candidate Set opsin gene class from figure 1: ciliary (brown), UV (violet), blue (blue), long wavelength (green) and outgroups (dark grey).

non-visual opsin of *P. nigroadumbratus* was grouped with the parapinopsin-like *B. mori*, nested within the invertebrate ciliary opsin clade, which formed a sister group to the remaining rhabdomeric visual opsins. The predicted *uvop* transcripts of both focal surface dytiscid species (*A. bistrigatus* and *P. nigroadumbratus*) formed a distal monophyletic group that was most closely related to UV opsins of *T. marmoratus*, the only other dytiscid beetle in the matrix. For long-wavelength opsin the same family level grouping for Dytiscidae was apparent. The only subterranean dystiscid to express an opsin (*L. palmulaoides*) formed a monophyletic group with the predicted *lwop* transcript of the other tribal member of Bidessini, *A. bistrigatus*. The *lwop* transcript of *P. nigroadumbratus* (Hydroporini) formed the sister branch to the Bidessini clade and *T. marmoratus* (Aciliini) was the next basal *lwop* branch. The dytiscid long-wavelength opsin clade was a sister group to bee and butterfly clades; however, the *T. castaneum* copy was recovered as sister lineages to all other long-wavelength opsins.

4.5. Comparative tests of evolutionary rates between surface and subterranean beetles

Considering that a blind (eye-less) species in an aphotic niche has expressed a visual opsin, we undertook analyses of evolutionary rates on *lwop*. It is recommended that codon based likelihood estimates of \( dN/dS \) are best undertaken on single gene products [45], so we removed all non-*lwop* opsin copies from their respective matrices for comparative analyses. Pairwise analyses were only performed on beetles in order to contrast surface and subterranean lineages. For phylogenetically informed analyses, the number and phylogenetic distance of taxa within an alignment matrix can influence results, hence we ran two datasets: one containing dytiscid beetles only (\( n = 4 \)), and a second matrix with a wider variety of insect outgroups including the red flour beetle, two bee and three butterfly *lwop* copies (\( n = 10 \)).

4.5.1. Pairwise estimates of \( dN/dS \)

Table 2 displays pairwise results that compare the stygobitic *L. palmulaoides* with all other surface dytiscids, followed by comparisons among the surface taxa. All *P*-distance comparisons of *lwop* rejected
neutral evolution, showed strong support for purifying selection and showed no evidence of positive selection. \textit{L.} sequences of the two \textit{Bidessini} species were very similar, with a nucleotide \textit{P}-distance of only 0.006, and amino acid products were identical.

### 4.5.2. Phylogenetically informed likelihood estimates of \( \text{dN}/\text{dS} \)

For the beetle-only matrix, AIC model tests selected 010123 for the nucleotide model and F81 (000000) for the codon model. The global \( \text{dN}/\text{dS} \) ratio (total evidence across the entire matrix and tree) was \( \omega = 0.0239475 \); however, an LRT comparing global \( \omega(H_0) \) with local \( \omega(H_1) \) rates favoured the alternate hypothesis \( (\chi^2 = 2660.59, p < 0.001) \). The expanded insect dataset \( (n = 10) \) was analysed with AIC determined nucleotide \( (012032) \) and codon \( (010110) \) models. There was no recombination in the alignment and the global \( \text{dN}/\text{dS} \) ratio \( \omega = 0.0463382 \) was rejected in favour of local rates \( (\chi^2 = 36.0746, p = 0.003) \). Based on these rejections of the shared global \( \omega \) ratio, we then explored a range of analyses contrasting branches within the tree.

### 4.5.3. Variation in \( \text{dN}/\text{dS} \) among branches

We compared evolutionary rates among various terminal branches of the tree using LRTs, exploring different local rate variations: (i) no site-to-site variation, (ii) \( \text{dN} \) rate variation only and (iii) independent \( \text{dN} \) and \( \text{dS} \) rate variation. We tested three alternative sets of branches against the rest of the tree: we first compared the subterranean \textit{L. palmulaoides} terminal branches \( (n = 3) \), then the surface beetle terminal branches \( (n = 3) \), and finally all terminal branches against all internal branches \( (n = 4) \). None of these comparisons showed significantly different rates of variation among select branches in the tree when tested against a global rate (electronic supplementary material, table S4). Similarly, the expanded insect dataset did not result in any significant LRT comparisons (electronic supplementary material, table S4).

### Table 2.

<table>
<thead>
<tr>
<th>photic niche</th>
<th>species</th>
<th>( P )-dist</th>
<th>( P )-dist</th>
<th>( H_0 ) neutral evolution ( \text{dN} = \text{dS} )</th>
<th>( H_1 ) non-neutral evolution 2-tailed ( \text{dN} \neq \text{dS} )</th>
<th>( H_2 ) purifying selection 1-tailed ( \text{dN} &lt; \text{dS} )</th>
<th>( H_3 ) positive selection 1-tailed ( \text{dN} &gt; \text{dS} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>surface</td>
<td>\textit{Allodessus bistrigatus}</td>
<td>0.0</td>
<td>0.006</td>
<td>( Z = -2.276, p = 0.025 )</td>
<td>( Z = 2.164, p = 0.016 )</td>
<td>( Z = -2.214, p = 1.0 )</td>
<td></td>
</tr>
<tr>
<td>surface</td>
<td>\textit{Limbodessus palmulaoides}</td>
<td>±0.0</td>
<td>±0.002</td>
<td>( Z = -15.902, p &lt; 0.001 )</td>
<td>( Z = 16.319, p &lt; 0.001 )</td>
<td>( Z = -16.026, p = 1.0 )</td>
<td></td>
</tr>
<tr>
<td>surface</td>
<td>\textit{Paroster nigroadumbratus}</td>
<td>0.119</td>
<td>0.239</td>
<td>( Z = -16.964, p &lt; 0.001 )</td>
<td>( Z = 16.298, p &lt; 0.001 )</td>
<td>( Z = -16.549, p = 1.0 )</td>
<td></td>
</tr>
<tr>
<td>surface</td>
<td>\textit{Thermonectus marmoratus}</td>
<td>0.128</td>
<td>0.239</td>
<td>( Z = -15.945, p &lt; 0.001 )</td>
<td>( Z = 16.388, p &lt; 0.001 )</td>
<td>( Z = -16.187, p = 1.0 )</td>
<td></td>
</tr>
<tr>
<td>surface</td>
<td>\textit{Allodessus bistrigatus}</td>
<td>0.119</td>
<td>0.239</td>
<td>( Z = -17.538, p &lt; 0.001 )</td>
<td>( Z = 16.964, p &lt; 0.001 )</td>
<td>( Z = -16.874, p = 1.0 )</td>
<td></td>
</tr>
<tr>
<td>surface</td>
<td>\textit{Thermonectus marmoratus}</td>
<td>0.143</td>
<td>0.235</td>
<td>( Z = -14.190, p &lt; 0.001 )</td>
<td>( Z = 13.954, p &lt; 0.001 )</td>
<td>( Z = -13.391, p = 1.0 )</td>
<td></td>
</tr>
</tbody>
</table>
4.5.4. Variation in $d_{N}/d_{S}$ at the site-specific (codon) level

Site-specific methods reiterated the results from pairwise Z-tests. A minimum of 10 sequences is recommended for these analyses [43], thus we only present results from the expanded ‘insect’ dataset. The $\text{lwop}$ matrix contained 328 codons, for which there were 498 distinct sites. Approximately one-half to two-thirds were estimated as undergoing purifying selection, depending on the analytical procedure: SLACs identified 153 codons (47%); fixed effects likelihood analysis identified 214 codons (65%). We also used FEL to compare the dytiscid beetle sub-tree in isolation to the remainder of the branches in the tree, which recovered 147 codons (45%), followed by the surface ($A. \text{bistrigatus}$) and subterranean ($L. \text{palmulaoides}$) bidessine species in isolation, which yielded only three codons (less than 1%). None of these perturbations identified any codons as undergoing positive selection. Thus, the subterranean and surface bidessines indicate very minimal change in the long-wavelength opsin expressed in the aphytic environment.

5. Discussion

Our initial aim was to qualitatively identify transcribed opsin genes in organisms occupying radically different photic niches. Under a neutral evolution theory, we expected detection of opsin transcripts in surface lineages, but a lack of transcription in subterranean lineages. Such data then provide grounds for subsequent comparative exploration of genomic DNA to assess whether non-adaptive (neutral) evolution has rendered these genes non-functional and contributed to the regressive evolution of eyes in stygofauna.

Our results show that in three independently evolved subterranean dytiscid beetle species nearly all opsin photoreceptor genes lacked evidence of transcription (except $\text{lwop}$ in $L. \text{palmulaoides}$), whereas the two surface-water species showed evidence of expression for a full suite of insect visual and non-visual opsins (except blue opsin). This result is suggestive of parallel molecular evolution of a loss of photoreceptor function, more specifically the ‘neutral’ regressive evolution of eye genes. An alternative explanation for a lack of opsin gene transcription in the subterranean species is that opsins were transcribed, but at extremely low levels that were insufficient to allow detection (i.e. a problem of read coverage). However, transcriptome sequence reads of the subterranean species were as high (or higher) than that of the surface species (electronic supplementary material, table S3), and greater than that used in other similar studies of subterranean taxa that identified transcription of opsin genes [46].

The $\text{lwop}$ orthologue transcribed by the subterranean dytiscid $L. \text{palmulaoides}$ did not exhibit rates of mutation consistent with neutral evolution ($d_{N} = d_{S}$); rather, there was strong evidence for purifying selection ($d_{N} < d_{S}$). The translated opsin product seems to be structurally functional as it is almost identical in sequence to orthologues from surface species. A similar result was found for the cave-dwelling carrion beetle $P. \text{hirtus}$ (which possesses a residual eye lens), for which transcriptome evidence suggests a reduced, but apparently functional, visual system [46].

Opsin apo-proteins bind with chromophores to form photopigment molecules that are responsible for the transduction of photons of light into neurophysiological signals, and there are visual and non-visual opsins. Rhabdomeric opsins of diurnally active insect eyes are typically sensitive to UV, blue and long wavelengths, which enables tri-chromatic colour vision; a number of non-visual ciliary opsins have been identified, but their function remains enigmatic. Our study has identified two visual opsins in surface-water dytiscid beetles orthologous to insect UV and long-wavelength classes and one non-visual ciliary opsin.

We did not identify a blue opsin orthologue at all in dytiscids, but note that its loss has been documented for the red flour beetle [47], and the only other publicly available sequence data for a dytiscid beetle’s opsins is from larvae of the sunburst diving beetle, which also lacks blue sensitivity [48,49]. Maintenance of blue sensitivity is likely to be dependent on the photic niche of the focal species, and electroretinogram studies suggest day-active predatory lady beetles do possess blue-sensitive (420 nm) photopigments [50].

The non-visual opsin we identified in surface-water dytiscids shows similarity to honeybee $\text{pteropsin}$ [51] and most closely resembles a predicted $\text{parapinopsin}$-like gene from silk worms; $\text{parapinopsin}$ is a non-visual opsin expressed in the pineal and parietal organs of vertebrates [52,53]. However, the role of non-visual ciliary-type opsins in insects is not comprehensively understood.
5.1. Opsin transcription in an aphytic subterranean habitat: pleiotropy or incipient pseudogene?

Our finding of expression of the visual \textit{lwop} by the subterranean \textit{L. palmulaoides} is counterintuitive. As indicated in the opsin phylogeny branch lengths and P-distances, there is minimal sequence difference between \textit{lwop} from \textit{A. bistrigatus} and \textit{L. palmulaoides}, there is no evidence for neutral evolution, there are no aberrant stop codons indicative of pseudogenes and evolutionary rates appear to be comparable to both close and distant phylogenetic insect relatives. So are we observing a precursory phase of pseudogene development? This would seem plausible if aphytic lineages have only very recently diverged from photic (surface) ancestors, as in some cavefish [12]. However, the subterranean beetles are proposed to have colonized calcrete aquifers during the Plio-Pleistocene aridification of Australia [22], with a major radiation of these subterranean beetles estimated between 3 and 8 Ma, based on uncalibrated molecular-clock analyses [23].

While the transcription of non-functional opsins over considerable evolutionary periods is possible (e.g. more than 100 Ma [54]), there are other examples whereby cave-adapted animals maintain seemingly functional visual opsins. Arthropod studies comparing cave- and surface-dwelling crayfish species [13] and amphipod populations [14] imply that there has been no loss of opsin function in the aphytic environment. Both studies conclude that this constitutes evidence for a pleiotropic role for the respective visual opsin proteins, but supporting empirical evidence for pleiotropy is limited. It should be noted that: (i) cave crayfish are cephalically blind but possess non-visual \textit{caudal} photoreceptors [55] that may be the primary means of photoreception [56], whereas the cave amphipods possess reduced eyes with far fewer ommatidia [14,57, p. 111]); and (ii) neither of these aforementioned cave environments physically isolate the organisms from their respective surface relatives, which is the case for the calcrete aquifers that house the Australian dytiscid beetle stygofauna [17,35,36], and an important distinction between our study system and most other explorations of regressive evolution in eye genes.

Among insects there is evidence of functional visual opsin expression in the larval and adult brains of butterflies and bees: silkworm \textit{lwop} [58], hawkmoth \textit{lwop} [59] and bumblebee \textit{uvop} [33]. Lampel \textit{et al.} [59] speculated on the circadian role of invertebrate non-visual opsins, and in at least one vertebrate, blind cavefish, point mutations (stop codons) in both \textit{melanopsin} and \textit{teleost multiple tissue opsin} are implicated in the de-activation of diurnal circadian rhythm [32]. The retention (‘recycling’) of larval-stage visual organs in adults of some endopterygote insects are also implicated in deep-brain circadian entrainment (notably \textit{Drosophila}, \textit{Tribolium} and \textit{Thermonectus}—reviewed in [60,61]), suggesting ontogenesis of these insect visual systems is complex and deserving of considerable attention.

The expression of opsin in the early developmental phases of \textit{Astyanax} cave-dwelling fish with degenerate eyes has been used as evidence for a pleiotropic role [62]. Juvenile cavefish continue to express an apparently functional red-opsin protein (cf. surface lineages [63]) for the first 3–6 days of development, which would be prior to the emergence of a functional eye in surface lineages. Langecker \textit{et al.} [62] postulated that opsin might play a structurally important role in the morphogenesis of normal eye development, but argued that eye regression must, therefore, be due to malfunctions in regulatory genes in \textit{Astyanax} cavefish.

More recent evolutionary studies on an alternative cavefish system, the amblyopsids [12], suggest that retention of a structurally sound opsin in blind morphs is a function of evolutionary time relative to colonization of the aphytic niche: amblyopsid species that have persisted for considerably longer periods in caves (up to 10 Ma) show evidence of a loss of selective constraint on opsins and subsequent pseudogene development [12]. We do not currently have a wide enough sampling of taxa in our system to be confident in linking the comparative age of subterranean colonization by \textit{L. palmulaoides}, relative to other stygobitic species, as a factor contributing to the retention of a seemingly functional \textit{lwop} protein.

6. Conclusion

This study presents a comparative assessment of the expression of visual and non-visual opsin proteins in surface (photic) and subterranean (aphotic) habitats of dytiscid water beetles. The blind subterranean species \textit{L. palmulaoides} expressed a structurally functional copy of \textit{lwop} that is indistinguishable in amino acid sequence from its surface relative \textit{A. bistrigatus}, and these affinities are supported by fully resolved phylogenetic trees. We cannot rule out a pleiotropic function for \textit{lwop} in dytiscid beetles, but note that an incipient phase of pseudogene development is also a plausible explanation for our finding. In-solution hybridization capture methods, which use RNA-baits to target nuclear DNA, carried out on a much wider sampling of species from this system will permit a more considered assessment of potential pseudogene development among subterranean species. This project embodies the power of
genomic-level investigations of evolutionary ecology, whereby the effects of environment and genetic inheritance can be dovetailed to explore phenotypic development and the associated physiology via gene transcription. Our finding of a general lack of opsin expression among three independently evolved subterranean diving beetle lineages suggests a parallel loss of opsin expression in an aphytic environment, which we argue is indicative of ‘neutral’ regressive evolution.

Data accessibility. Sequences of opsin orthologues are deposited in GenBank under accession numbers KP219380–KP219386, and nexus file alignment data are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.0dzq8.

Acknowledgements. Sincere thanks to C. Watts for surface-water beetle collections and identifications; also to R. Leis and J. Walcock for subterranean collections. We are grateful to eResearch SA for bioinformatic support, and to general staff of the University of Adelaide, South Australian Museum and Western Australian Museum for logistical assistance. H. Wilkens, M. Guzik and two anonymous reviewers provided insightful comments on previous versions of the manuscript.

Author contributions. S.M.T. participated in the molecular laboratory work, designed and undertook all aspects of bioinformatic analyses (transcriptome assembly, orthologue search, phylogenetic reconstruction, tests of selection) and drafted the manuscript; S.J.B.C. conceived and coordinated the study, collected field samples, assisted with design of bioinformatics, and helped draft the manuscript; K.M.S. undertook molecular laboratory work; T.B. assisted with bioinformatics (assembly and orthologue search); J.H. assisted with orthologue searches; W.E.H. conceived the study and collected field samples; A.D.A. conceived and coordinated the study. All authors commented on the manuscript and gave final approval for publication.

Funding statement. Research was supported by Australian Research Council grant DP120102132 awarded to S.J.B.C., W.F.H. and A.D.A.

Competing interests. We have no competing interests.

References
