INVITED REVIEW

Evolution in caves: Darwin’s ‘wrecks of ancient life’ in the molecular era

CARLOS JUAN,* MICHELLE T. GUZIK,† DAMIÀ JAUME‡ and STEVEN J. B. COOPER†§
*Departament de Biologia, Universitat de les Illes Balears, 07122 Palma de Mallorca, (Balearic Islands) Spain, †Australian Centre for Evolutionary Biology and Biodiversity and School of Earth and Environmental Science, The University of Adelaide, Adelaide, SA 5005, Australia, ‡IMEDEA (CSIC-UIB), Mediterranean Institute for Advanced Studies, c/Miquel Marquès 21, 07190 Esporles (Balearic Islands), Spain, §Evolutionary Biology Unit, South Australian Museum, North Terrace, Adelaide, SA 5000, Australia

Abstract

Cave animals have historically attracted the attention of evolutionary biologists because of their bizarre ‘regressive’ characters and convergent evolution. However, understanding of their biogeographic and evolutionary history, including mechanisms of speciation, has remained elusive. In the last decade, molecular data have been obtained for subterranean taxa and their surface relatives, which have allowed some of the classical debates on the evolution of cave fauna to be revisited. Here, we review some of the major studies, focusing on the contribution of phylogeography in the following areas: biogeographic history and the relative roles of dispersal and vicariance, colonization history, cryptic species diversity and modes of speciation of cave animals. We further consider the limitations of current research and prospects for the future. Phylogeographic studies have confirmed that cave species are often cryptic, with highly restricted distributions, but have also shown that their divergence and potential speciation may occur despite the presence of gene flow from surface populations. Significantly, phylogeographic studies have provided evidence for speciation and adaptive evolution within the confines of cave environments, questioning the assumption that cave species evolved directly from surface ancestors. Recent technical developments involving ‘next generation’ DNA sequencing and theoretical developments in coalescent and population modelling are likely to revolutionize the field further, particularly in the study of speciation and the genetic basis of adaptation and convergent evolution within subterranean habitats. In summary, phylogeographic studies have provided an unprecedented insight into the evolution of these unique fauna, and the future of the field should be inspiring and data rich.

Keywords: cave animals, cryptic species, phylogeography, speciation, subterranean, vicariance and dispersal

Received 8 April 2010; revision received 7 June 2010; accepted 10 June 2010

Introduction

Darwin (1859) identified cave animals, which often show peculiar convergent ‘regressive’ traits, such as absence of eyes and lack of body pigment, as ‘wrecks of ancient life’. Indeed, the singularity of subterranean fauna has stimulated the imagination of several generations of biologists since the discovery of the first cave species, the salamander Proteus anguinus Laurenti, 1768 (Barr 1967; Poulson & White 1969; Culver 1970). Cave ecosystems have been considered natural ecological and evolutionary laboratories because of the relative simplicity of their communities and the temporal and spatial isolation of their biota (Poulson & White 1969). Thus, cave animals can help in our understanding of
evolutionary processes and historical factors related to biogeography and speciation in a similar manner to that of isolated oceanic island biotas (Cooper et al. 2007; Page et al. 2008). More than 7000 aquatic subterranean species (stygobionts) are known (Gibert & Deharveng 2002), whereas estimates of the number of terrestrial cave taxa (troglobionts) point to a minimum of 21 000 worldwide (Juberthie & Decu 1994). These figures are likely to be an underestimation since caves and other subterranean environments are poorly studied, and these habitats in many regions of the world still remain to be explored (Gibert & Deharveng 2002).

Historically, many hypotheses concerning the evolution of cave animals have proven to be controversial, and some are still a matter of debate. These include the climate relict vs. ecological/parapatric modes of speciation by adaptive shift (Howarth 1973; Holsinger 2000; Rivera et al. 2002; Wessel et al. 2007), the role played by dispersal vs. vicariance to explain subterranean biogeographic patterns (see Porter 2007 for a review) and selective (Yamamoto & Jeffery 2000; Jeffery 2005; Romero & Green 2005) vs. neutral (Kosswig 1960; Culver & Wilkens 2000; Leys et al. 2005) hypotheses to explain regressive evolution (see Jeffery 2009 and Wilkens 2010 for recent reviews in cave fish). Finally, the irreversibility of ecomorphological adaptations to the subterranean (hypogean) habitat and the possibility of recolonization of surface (epigean) habitats by hypogean lineages have been recently debated as well (Humphreys 2000; Collin & Miglietta 2008; Volschenk & Prendini 2008; Prendini et al. 2010).

In the last decade, a wealth of molecular data has been gathered from both terrestrial and aquatic subterranean taxa (see Table 1 for a list of publications). These data have allowed the testing of hypotheses on biogeography and modes of subterranean evolution within explicit phylogenetic frameworks, distinguishing the role played by ecological adaptations and by historical factors in the evolution of cave faunas (Arnedo et al. 2007; Juan & Emerson 2010). Here, we review the major contributions that molecular data have made in the field of subterranean biology, describing recent case studies from the literature. We also assess the limitations of current studies and discuss future prospects and directions in this emerging area. We focus on studies in the fields of biogeography/phylogeography and speciation, complementing a recent review of the neutral vs. selection hypotheses of regressive evolution (Wilkens 2010).

The biogeographic history of subterranean animals: the role of dispersal and vicariance

There has been a long and recurrent debate on the relative roles played by vicariance and dispersal in the biogeography of subterranean taxa (see Culver et al. 2009 for a historical review). Traditionally, the extremely localized distributions of many subterranean taxa have been explained by their highly reduced dispersal capabilities because of intrinsic (low vagility, high specialization, small size) and extrinsic (habitat fragmentation) factors (Crouau-Roy 1989; Holsinger 1991; Gibert et al. 1994). The relative contribution of dispersal and vicariance varies among geographical regions and between terrestrial and aquatic subterranean taxa, the latter usually showing broader ranges because of the higher connectivity between subterranean aquifers compared to karstic terrestrial habitats (Porter 2007; Culver et al. 2009).

A combination of dispersal and vicariance events is suggested to have influenced the biogeographic history of peri-Thyrrenian stenasellid isopods (Ketmaier et al. 2003). The detachment of the Corsico-Sardinian microplate from the Pyrenees explains an old divergence in these isopods (between Stenasellus racovitzai and the mainland Stenasellus virei), while dispersal accounts for the unexpected low genetic divergence found between mitochondrial cytochrome oxidase subunit 1 (cox1) sequences of S. racovitzai lineages of Tuscany and Corsica (with an estimated divergence time of 2–0.5 million years ago (Ma), thus long after than the separation of the island from the Italian Peninsula). These results are based on a calibration of the molecular clock using two different geological events; the split of Corsica–Sardinia from the Pyrenees 29 Ma, and the separation between Sardinia and Corsica (9 Ma).

The extreme disjunct global distributions displayed by many stygobiont crustaceans are thought to result from plate tectonic vicariant events that subdivided widely distributed ancestral populations once living along the margins of late Mesozoic seas (the full Tethyan track hypothesis, see Fig. 1, Stock 1993; Jaime & Christenson 2001). To date, this hypothesis has not been well explored using molecular methods. However, Page et al. (2008) recently used mitochondrial and nuclear sequences to study the endemic Western Australian subterranean shrimp genus Stygiocaris (Atyidae) and its relationship with other Australian and North Atlantic relatives. These authors conclude that the closest relative to the West Australian Stygiocaris is the Mexican cave genus Typhlatya. Despite the presumed dispersal capacity of these shrimps during their larval stages, molecular data suggested that the two sister groups derived from an ancient vicariant event, namely the cessation of migration in populations of a coastal marine ancestor about 19 Ma by the closure of the Tethys Seaway following tectonic plate movements. Future molecular analyses should confirm if the extreme disjunct distributions observed in other
Table 1 A list of the main studies on biogeographic and colonization histories, cryptic species diversity and modes of speciation published in the last decade using DNA molecular markers in cave-dwelling fauna. Mitochondrial and nuclear markers are indicated for each study.

<table>
<thead>
<tr>
<th>Taxonomic group</th>
<th>Mitochondrial</th>
<th>Nuclear</th>
<th>Geographical region</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Markers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Annelids</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polychaetes</td>
<td>–</td>
<td>SSU, LSU</td>
<td>Dinaric karst</td>
<td>Kupriyanova et al. 2009</td>
</tr>
<tr>
<td>Crickets</td>
<td>cox1, rnl</td>
<td>–</td>
<td>W Mediterranean</td>
<td>Allegrucci et al. 2005</td>
</tr>
<tr>
<td>Beetles</td>
<td>cox1</td>
<td>–</td>
<td>W Mediterranean</td>
<td>Caccone &amp; Sbordoni 2001</td>
</tr>
<tr>
<td></td>
<td>cox1, rnl, nad1</td>
<td>–</td>
<td>W Australia</td>
<td>Cooper et al. 2002</td>
</tr>
<tr>
<td></td>
<td>cox1, rnl, nad1</td>
<td>–</td>
<td>W Australia</td>
<td>Leys et al. 2003</td>
</tr>
<tr>
<td></td>
<td>cox1, rnl, nad1</td>
<td>–</td>
<td>W Australia</td>
<td>Balke et al. 2004</td>
</tr>
<tr>
<td></td>
<td>cox1, cox2</td>
<td>ITS-2</td>
<td>Macaronesia</td>
<td>Contreras-Diaz et al. 2007</td>
</tr>
<tr>
<td></td>
<td>cox1, rnl, nad1</td>
<td>SSU, Histone H3</td>
<td>China</td>
<td>Wewalka et al. 2007</td>
</tr>
<tr>
<td></td>
<td>cox1, rnl</td>
<td>–</td>
<td>NW Australia</td>
<td>Leys &amp; Watts 2008</td>
</tr>
<tr>
<td></td>
<td>cox1, rnl, nad1</td>
<td>SSU, LSU</td>
<td>W Mediterranean</td>
<td>Faille et al. 2010</td>
</tr>
<tr>
<td></td>
<td>cox1, rnl, nad1</td>
<td>SSU, LSU</td>
<td>W Mediterranean</td>
<td>Ribera et al. 2010</td>
</tr>
<tr>
<td></td>
<td>cox1, rnl, nad1</td>
<td>–</td>
<td>W Australia</td>
<td>Guzik et al. 2009</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>W Mediterranean</td>
<td>Caccone &amp; Sbordoni 2001</td>
</tr>
<tr>
<td><strong>Insects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Arachnids</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spiders</td>
<td>cox1</td>
<td>–</td>
<td>Texas, US</td>
<td>Paquin &amp; Hedin 2004</td>
</tr>
<tr>
<td></td>
<td>cox1, rnl</td>
<td>–</td>
<td>Canary Is.</td>
<td>Arnedo et al. 2007</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>SE United States</td>
<td>Hedin &amp; Thomas 2010</td>
</tr>
<tr>
<td>Pseudoscorpions</td>
<td>cox1</td>
<td>–</td>
<td>S Australia</td>
<td>Moulds et al. 2007</td>
</tr>
<tr>
<td><strong>Molluscs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gastropods</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bivalves</td>
<td>cox1, rnl</td>
<td>–</td>
<td>Dinaric karst</td>
<td>Stepien et al. 2001</td>
</tr>
<tr>
<td></td>
<td>rnl</td>
<td>–</td>
<td>Philippines</td>
<td>Kano &amp; Kase 2004</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>ITS</td>
<td>Borneo</td>
<td>Schilthuizen et al. 2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>France</td>
<td>Bichain et al. 2007</td>
</tr>
<tr>
<td><strong>Crustaceans</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amphipods</td>
<td>cox1</td>
<td>LSU</td>
<td>Central Europe</td>
<td>Lefebure et al. 2006</td>
</tr>
<tr>
<td></td>
<td>cox1, rnl</td>
<td>LSU</td>
<td>France</td>
<td>Lefebure et al. 2007</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>NW Australia</td>
<td>Finston et al. 2007</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>W Australia</td>
<td>Cooper et al. 2007</td>
</tr>
<tr>
<td></td>
<td>cox1, cox2</td>
<td>Histone H3</td>
<td>La Palma (Canary Is.)</td>
<td>Villacorta et al. 2008</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>Australia</td>
<td>Murphy et al. 2009</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>ITS-1</td>
<td>Virginia, US</td>
<td>Carlini et al. 2009</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>LSU</td>
<td>Europe</td>
<td>Trontelj et al. 2009</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>W Australia</td>
<td>Bradford et al. 2010</td>
</tr>
<tr>
<td>Isopods</td>
<td>cox1</td>
<td>–</td>
<td>Hawaii Is.</td>
<td>Rivera et al. 2002</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>W Mediterranean</td>
<td>Ketmaier et al. 2003</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>RAPDs</td>
<td>NW Balkans</td>
<td>Verovnik et al. 2003</td>
</tr>
<tr>
<td></td>
<td>rna5, rnl</td>
<td>–</td>
<td>W Mediterranean</td>
<td>Baratti et al. 2004</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>NW Balkans</td>
<td>Verovnik et al. 2004</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>W Australia</td>
<td>Cooper et al. 2008</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>W Australia</td>
<td>Finston et al. 2009</td>
</tr>
<tr>
<td>Mysids</td>
<td>cox1</td>
<td>–</td>
<td>NW Mediterranean</td>
<td>Lejeusne &amp; Chevaldonné 2006</td>
</tr>
<tr>
<td>Decapods</td>
<td>rnl</td>
<td>–</td>
<td>SE United States</td>
<td>Buhay &amp; Crandall 2005</td>
</tr>
<tr>
<td></td>
<td>rnl</td>
<td>–</td>
<td>E United States</td>
<td>Finlay et al. 2006</td>
</tr>
<tr>
<td></td>
<td>cox1, rnl</td>
<td>LSU</td>
<td>France, W Balkans</td>
<td>Zaks et al. 2007</td>
</tr>
</tbody>
</table>

© 2010 Blackwell Publishing Ltd
anchialine (caves flooded by a mixture of marine water and terrestrial fresh groundwater lacking surface connections to the sea) groups of presumed Tethyan distribution, such as remipedes and some ostracod, copepod and thermosbaenacean crustaceans, have parallel genetic divergences dating back to the Tethys sea. However, it can be anticipated that a combination of mitochondrial and conserved nuclear DNA sequences in conjunction with different calibration points based on geological datings will be needed to robustly test this hypothesis.

Vicariance followed by secondary contact also seem to have operated over a more recent timescale in *Stygicaris*, with evidence provided for an east–west separation of the two currently recognized species, with the occurrence of a zone of sympatry in the northern portion of the Cape Range peninsula in Western Australia. This pattern has been related to isolation by the orogenic upraising of the peninsula followed by limited secondary contact during low sea level stands (Page et al. 2008). One case demonstrating the occurrence of limited dispersal and coexistence of closely related subterranean taxa was recently revealed by Santos (2006) and Craft et al. (2008), who found low levels of gene flow and a high level of genetic differentiation in populations of the endemic Hawaiian anchialine shrimp *Halocaridina rubra* (Atyidae) on and between islands of the archipelago. Nevertheless, cases are known among anchialine stygobionts with a broad distribution where population structuring is virtually nonexistent, suggesting the occurrence of panmixia mediated by marine dispersal. This is the case, for example, in the endemic Hawaiian alpheid shrimp *Metabetaeus lohena* (Russ et al. 2010), with populations on O‘Ahu, Maui and Hawaii; or in the neritilid gastropod *Neritilia granosa*, with populations set 200 km apart in the Philippines (Kano & Kase 2004).

Marine transgressions/regressions have been proposed to generate endemic inland freshwater species derived from a common marine ancestor (Notenboom 1991; Holsinger 1994, 2000). The origin of many stygobiontic taxa (stygofauna) whose distribution falls within areas formerly covered by ancient seas has been explained by the stranding of populations of a marine ancestor in inland groundwaters followed by divergence and speciation as the sea retreats (Zakšek et al. 2007). However, some cases of direct surface freshwater derivation in presumed lineages of marine origin have

---

Table 1 (Continued)

<table>
<thead>
<tr>
<th>Taxonomic group</th>
<th>Mitochondrial</th>
<th>Nuclear</th>
<th>Geographical region</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anchialine</td>
<td>cox1, rrsS, rrlL</td>
<td>Histone H3, GAPDH</td>
<td>Tennessee, Alabama US</td>
<td>Buhay et al. 2007</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>Hawaii Is.</td>
<td>Santos 2006</td>
</tr>
<tr>
<td></td>
<td>cox1, rrlL</td>
<td>–</td>
<td>Australia &amp; Indo-Pacific</td>
<td>Page et al. 2007</td>
</tr>
<tr>
<td></td>
<td>cox1, rrlL</td>
<td>Histone H3</td>
<td>Australia &amp; Indo-Pacific</td>
<td>Page et al. 2008</td>
</tr>
<tr>
<td></td>
<td>cox1, cob, rrlL</td>
<td>–</td>
<td>Caribbean &amp; W Atlantic</td>
<td>Hunter et al. 2008</td>
</tr>
<tr>
<td></td>
<td>cox1, rrlL</td>
<td>ITS-2</td>
<td>W Balkans</td>
<td>Zakšek et al. 2009</td>
</tr>
<tr>
<td></td>
<td>cox1</td>
<td>–</td>
<td>Hawaii Is.</td>
<td>Russ et al. 2010</td>
</tr>
<tr>
<td>Bathynellacians</td>
<td>cox1</td>
<td>–</td>
<td>W Australia</td>
<td>Guzik et al. 2008</td>
</tr>
</tbody>
</table>

| Fish            | rrsL, control region | RAPDs | Thailand | Borowsky & Mertz 2001 |
|                 | –                   | –      | –        | –                     |
|                 | cob                 | Microsatellites | Mexico | Streeker et al. 2003 |
|                 | cob, rrlL, rrsS     | –      | Mexico, Belize, Guatemala | Streeker et al. 2004 |
|                 | –                   | –      | SE Arabian Peninsula | Li et al. 2008 |
| Salamanders     | cob                 | Allozymes | Texas, US | Wiens et al. 2003 |
|                 | cob, nad4           | –      | E North America | Bonett & Chippindale 2004 |
|                 | cob, rrsS, control region | – | Balkans | Gorički & Trontelj 2006 |
|                 | rrsS, cob           | RAG-1  | Tennessee, US | Niemiller et al. 2008 |

RAPD, random amplified polymorphic DNA; RAG-1, recombination activating gene 1; GAPDH, glyceraldehyde-3-phosphate dehydrogenase, ITS, internal transcribed spacer.
been revealed by molecular analyses. Mitochondrial and nuclear phylogenies have shown that three different lineages (France-Iberian Peninsula, Dinaric-Caucasian and Western Caucasus clades) of the South European atyid cave shrimp *Troglocaris* are related to three different surface ancestors. A freshwater common ancestor of Caucasian and Dinaric cave shrimps was dated at 6–11 Myr, contemporary with the brackish and freshwater phase of the Paratethys (Zaksˇek et al. 2007).

The same holds for two subterranean atyid genera from northern Australia (*Parisia*, *Pycnisia*), closely related to surface freshwater species of *Caridina* of the same region (Page et al. 2007, 2008).

In a comparative phylogeographic analysis, Trontelj et al. (2009) used mitochondrial and nuclear gene sequences from six freshwater subterranean species showing apparently broad ranges on the European mainland: *Niphargus virei* and *Niphargus rhenorhodenensis* (Amphipoda, Niphargidae), *Troglocaris anophthalmus*, *Troglocaris hercegovinensis* and *Speleocaris pretneri* (Decapoda, Atyidae) and *Proteus anguinus* (Amphibia, Caudata, Proteidae). They uncovered substantial levels of cryptic diversity within each of these species (see the section on cryptic species below), with most of the identified lineages having narrow distributions. These findings lend support to the hypothesis that stygofauna generally have very narrow distributions because of hydrographical isolation and their low dispersal abilities.

Several molecular studies have been carried out on the recently discovered and largely unknown subterranean stygofauna of the isolated terrestrial limestone aquifers (called calcrites) of inland Western Australia. This faunistic assemblage includes Dytiscidae (diving beetles) (Cooper et al. 2002; Leys et al. 2003; Leys & Watts 2008), Amphipoda (Cooper et al. 2007; Finston et al. 2007), Isopoda (Cooper et al. 2008; Finston et al. 2009) and Bathynellacea crustaceans (Guzik et al. 2008).

The molecular phylogenies obtained from 60 species of diving beetles belonging to the tribes Bidessini and Hydroporini (Coleoptera: Dytiscidae), together with relaxed molecular clock calibrations, showed that individual calcrite aquifers in the Yilgarn region of Western Australia contain distantly related lineages derived from surface ancestors that independently colonized the calcrites between 9 and 4 Ma (Leys et al. 2003). This pattern of relict lineages restricted to individual calcrites has been corroborated using *cox1* mitochondrial sequence data (with some uncertainties still remaining because of lack of known surface ancestors among other factors) from subterranean amphipods (Paramelitidae).
and Chiltoniidae; Cooper et al. 2007), oniscoidean isopods (Haloniscus, Cooper et al. 2008) and bathynellid syncarids (Guzik et al. 2008). Isolation of calcareous populations was deduced to have occurred at least since the Pliocene, coinciding with the occurrence of a major aridity phase that led to the intermittent drying up of surface waters (Byrne et al. 2008). Other recent molecular studies of subterranean amphipods and isopods carried out in the Pilbara region of Western Australia (Finston et al. 2007, 2009) show again that diversification among major lineages is consistent with the onset of arid conditions during the Pliocene (data deduced from cox1 mitochondrial phylogenies), a result consistent also with the estimated dates for central Western Australia (Yilgarn region). However, contrasting colonization patterns are deduced in the two groups: the phylogeny is compatible with a single common ancestor in the amphipod diversification, while multiple invasions explain the divergent lineages found in Pygolabis isopods (Finston et al. 2009).

The study of genetic variation in insects, in particular beetles, was initiated in the 1970’s by pioneering surveys using allozymic markers (e.g. Carmody et al. 1972; Laing et al. 1976; Cockley et al. 1977; Caccone & Sbordoni 1987; Crouau-Roy 1989). Molecular DNA phylogenies have also been obtained in the last few years for many terrestrial cave animals including beetles, isopods, crickets, pseudoscorpions, spiders and salamanders (see references in Table 1). Recent studies on terrestrial cave beetles suggest that the present distributions of some taxa tend to reflect those of ancestral surface ancestors and, presumably, could be caused by ancient vicariant events involving the establishment of barriers to dispersal in the subterranean domain (Faille et al. 2010; Ribera et al. 2010, see section on Speciation within subterranean habitats below and Fig. 2).

Colonization history of subterranean animals: insights from phylogeography

Understanding the evolutionary processes that lead to morphological specialization, and the genetics of speciation in cave animals require simple case-study models where recently evolved cave populations coexist with their presumed surface ancestors, as the early stages of the evolution of ancient subterranean taxa are difficult to reconstruct (Strecker et al. 2003). Phylogeographic frameworks and population genetic analyses can provide the basis for discriminating whether cave forms derive from a single colonization event followed by subterranean dispersal, or from multiple independent invasions of a fragmented subterranean habitat. This further allows discerning whether troglobiontic features (troglomorphies or ecomorphological adaptations to the hypogean habitat) have a single origin or appear by convergence and can also be informative on the dynamics of gene flow between surface and/or cave populations.

One of the best model study systems that comprise closely related surface and cave populations is the neotropical fish Astyanax fasciatus (syn. Astyanax mexicanus Filippi). This species is broadly distributed in Mexican surface waters and includes many depigmented and blind-fish populations in karstic caves on the northeast of the country (Avise & Selander 1972; Mitchell et al. 1977; Wilkens 1988; Espinasa & Borowsky 2001; Dowling et al. 2002; Strecker et al. 2003, 2004). The apparent independent evolution of cave populations and their co-existence and hybridization with nearby surface populations have made it an ideal model system to explore the genetic basis of regressive and adaptive evolution of cave animals (see Wilkens 2010 and references therein). Mitochondrial cytochrome b (cob) data clearly show that several of the present cave populations of this fish originated from an ancestral genetic stock different from the current surface populations and that a minimum of two independent invasions have occurred. The pattern is consistent with the extinction of surface populations, leaving behind relict underground populations because of Pleistocene climatic changes with subsequent recolonization and expansion of surface populations from the south (Dowling et al. 2002; Strecker et al. 2003, 2004). Microsatellite data also suggest that the occurrence of repeated population bottlenecks in cave populations and of introgressive hybridization between surface and underground populations is a common feature in Astyanax (Strecker et al. 2003).

The Dinaric karst region of the Balkan Peninsula in Europe, host to the world’s richest subterranean fauna, also offers considerable potential for phylogeographic and population genetic analyses of related surface and cave populations (Sket et al. 2004). Although the composition of this fauna is known fairly well, molecular analyses have only recently been undertaken (e.g. Verovnik et al. 2004; Zakšek et al. 2009) and are offering unique insights into the evolutionary histories of cave species and how they have been influenced by the complex hydrographic history of the region. Verovnik et al. (2004) used cox1 mitochondrial sequences from surface and troglobiontic populations of the crustacean isopod Asellus aquaticus to uncover its genetic variation and cave-surface relationships. In this study, samples grouped by hydrographic connection rather than by habitat, and at least three different colonizations of the caves were deduced to have occurred, probably after the ancestral
surface populations were already differentiated by vicariance in this extremely fragmented karst habitat (Verovnik et al. 2004).

The tropical snail genus Georissa (Hydrocenidae) in western Sabah, Malaysian Borneo, has a differentiated cave morph, which seems to be a direct derivative of the surface Georissa saulae (Schilthuizen et al. 2005). Connections between surface and subterranean populations have also been examined in the crayfish Cambarus tenebrosus (rrnL 16S rRNA mitochondrial gene) from the Cumberland Plateau of the southern Appalachians (Finlay et al. 2006). Phylogenetic and nested clade analyses demonstrated a lack of correspondence between habitat and phylogeny and no evidence for a recent invasion of the caves (Finlay et al. 2006).

Palmorchestia hypogaea, one of the few known terrestrial troglobiontic amphipods, occurs in lava tubes on the island of La Palma in the volcanic Canary archipelago. On the same island, a related surface form dwells in the leaf litter of the laurel forest (Stock & Martín 1988; Stock 1990). Villacorta et al. (2008) studied populations of the surface and cave forms across the island using cox1 and cox2 mitochondrial sequences, revealing strong phylogeographic structuring and evidence for multiple invasions of the lava tubes. In cave and surface populations of the freshwater amphipod Gammarus

---

**Fig. 2** Schematic representation of simplified phylogenies, approximate geographical ranges and diversification times of cave beetles of tribes Leptodirini (Leiodidae) and Trechini (Carabidae) in the Pyrenees. A: Aphaenops, G: Geaphaenops, H: Hydraphaenops. Numbers in the Leptodirini phylogeny refer to monophyletic clades that include species from different genera with particular geographical distributions. Phylogenetic reconstructions were based on mitochondrial (cox1, cob, rrnL, trnL and nad1) and nuclear (SSU and LSU) markers. Trees were reconstructed using Bayesian analyses and estimation of divergence times performed by a Bayesian relaxed phylogenetic approach implemented in BEAST v1.4.7 Drummond & Rambaut (2007) calibrated by standard mitochondrial substitution rates (Trechini) and biogeographical dating (Leptodirini). Photographs of the representative Pyrenean taxa Aphaenops alberti (Trechini) and Baronniesia delioti (Leptodirini) from Faille et al. (2010) and Fresneda et al. (2009) respectively. Figure modified from Faille et al. (2010) and Ribera et al. (2010).
minus, a similar pattern of spatial structuring, low genetic variation within caves or springs and relatively high divergence among sites with limited gene flow have been reported in the Appalachian Mountains (Carlini et al. 2009). As in the Palmorchestia study, population clustering is better explained by hydrological or geographic factors rather than by the type of habitat (i.e. whether cave or surface), and there is some evidence for the past occurrence of repeated population bottlenecks in cave populations (Carlini et al. 2009).

Assessing cryptic species diversity in cave environments using molecular data

Recently, phylogeographic analyses have uncovered considerable levels of cryptic species diversity within subterranean ecosystems. ‘Cryptic species’ usually refer to two or more taxa classified as a single species based on morphological similarity (Pfenninger & Schwenk 2007). These taxa can be a by-product of poor taxonomy (therefore ‘crypticism’ often disappears when adequate morphological features are considered) or be independent evolutionary lineages showing morphological stasis. Convergent evolution resulting from exposure to similar selection pressures, or lack thereof, within the lightless subterranean environment, can produce genetically divergent but cryptic morphological lineages (e.g. Culver et al. 1995; Arntzen & Sket 1997; Verovnik et al. 2003; Fiser et al. 2006; Hedin & Thomas 2010). In addition, convergence of form and adaptive modifications, such as loss of traits (Porter & Crandall 2003) or their gain (Jones et al. 1992), can render evolutionary relationships elusive since the shared morphology may not reflect shared phylogenetic history (Wiens et al. 2003).

Cryptic species have been described for surface crustaceans (e.g. King & Hanner 1998; Witt & Herbert 2000; Penton et al. 2004), and this is also an emerging issue for subterranean crustaceans (Finston et al. 2007; Buhay & Crandall 2009; Zakšek et al. 2009) and salamanders (Wiens et al. 2003).

Molecular data provide new insights into the levels of species diversity in cave systems. Historically, allozyme genetic divergence data permitted the discrimination of geographically distant cave populations and the possible occurrence of cryptic species (Cesaroni et al. 1981; Sbordoni et al. 1981; Kane et al. 1992; Allegrucci et al. 1997). Subsequent advances in sequencing technology have led to the use of DNA sequence data for the identification of cryptic species (e.g. King & Hanner 1998; Witt & Herbert 2000; Penton et al. 2004), particularly in subterranean crustaceans (Proudlove & Wood 2003; Finston et al. 2007; Buhay & Crandall 2009; Zakšek et al. 2009) and salamanders (Wiens et al. 2003).

The recent development of DNA barcoding thresholds for delineating species (16% divergence in cox1), genera and families based on crustacean sequence data has provided a new benchmark for species recognition (Lefebure et al. 2006). While these thresholds, coupled with phylogeographical analyses, have been useful in revealing the possible presence of cryptic species of subterranean animals (Trontelj et al. 2009), their generalized application poses several problems. First, recently evolved species cannot be discriminated, with several studies demonstrating the occurrence of lower levels of genetic divergence between morphologically distinct species (Leys et al. 2003; Guzik et al. 2008; Trontelj et al. 2009). Second, systematic categories such as genus and family are morphologically defined entities that do not necessarily reflect any clear timescale of evolution, particularly when different animal groups are compared. Finally, the high levels of saturation in DNA sequences at these higher taxonomic levels can pose additional problems. We conclude therefore that the use of molecular thresholds may have a place in delineating species taxonomic research; however, resolution above the species level is likely to be questionable.

Modes of speciation in subterranean animals

Two alternative models are proposed to explain speciation in cave faunas: the ‘climatic-relict’ and the ‘adaptive-shift’ hypotheses (Howarth 1973; Holsinger 2000; Rivera et al. 2002). The first model was proposed for continental temperate ecosystems (Holsinger 1988, 2000; Peck & Finston 1993) where invasion of the cave environment by surface species is coupled with an initial phase of gene flow between surface and cave populations. This phase is followed by speciation in strict allopatri (between cave and surface populations) after extinction of the epigean populations because of climatic fluctuations (e.g. glaciations or aridity). However, during the 1970s, a diverse cave fauna in the tropics was discovered along with some of the extent closely related surface relatives, leading to the proposal of the ‘adaptive-shift’ hypothesis (Howarth 1987; Roux & Danieilopol 1987; Desutter-Grandcolas & Grandcolas 1996). This model assumes the occurrence of active colonization of the subterranean environment accompanied by adaptive differentiation of surface and cave populations, reduction in gene flow and, eventually, parapatric speciation. A parapatric ‘adaptive-shift’ model is reliant upon divergent natural selection, while the allopatic model requires an accumulation of neutral or ‘regressive’ mutations, but may or may not involve adaptive shifts in relation to reproductive isolation. It is therefore the presence/absence of gene flow during divergence (and not selection) that is the key factor that allows discrimination between these two models (Niemiller et al. 2010).
An additional question is whether reproductive isolation can be reached without the involvement of an allopatric phase to complete the speciation process.

Molecular phylogenies could, in theory, help to test these two hypotheses as they allow reconstruction of sister relationships between cave and surface taxa (Rivera et al. 2002). In particular, phylogenetic contrasts coupled with the analysis of present geographical distributions (viz. surface and cave populations with allopatric distributions; relict subterranean species; or parapatric distributions of both surface and cave populations) can potentially distinguish between the above-mentioned hypotheses (Rivera et al. 2002; Leys et al. 2003). Additional tests have been proposed that correlate divergence times and branching patterns with geological/palaeo-climatic data (Leys et al. 2003). However, phylogenetic studies have some limitations, particularly when the divergence of cave lineages is very old and their related surface relatives became extinct, or have departed from the region after the advent of environmental changes. In addition, molecular clock estimates are subject to considerable errors, as few reliable rate calibrations exist for cave animals, whereas information is very limited on whether rates of molecular evolution increase or decrease when animals live underground. Finally, single or a few loci only, are often used to reconstruct phylogenetic relationships, with gene tree divergences times not necessarily equating to species divergence times.

In the subterranean diving beetles of the arid zone of Western Australia, the transition to the subterranean environment, especially in sympatric sister-species pairs, was coincident in time with a period of aridity that commenced approximately 10–5 Ma. This finding suggested that climate change (and thus a ‘climatic-relict’ allopatric speciation model) could be the main factor inducing speciation (Leys et al. 2003). On the contrary, in Hawaiian terrestrial isopods studied using coxl sequences, Rivera et al. (2002) showed that the sister relationship and parapatric distribution of cave and surface species of the genus Littorophiloscia could be explained by the occurrence of an adaptive shift from a marine supralittoral to a subterranean terrestrial lifestyle.

Two recent studies, the first focused on tropical land snails (Schilthuizen et al. 2005) and the other on North American salamanders (Niemiller et al. 2008), suggest the occurrence of speciation scenarios of ‘divergence with gene flow’ (Rice & Hostert 1993). An unnamed cave population of the land snail Georissa from Sabah (see above) showed a sister relationship to the surface Georissa saulae based on rrnL mitochondrial sequences and shell characters. The occurrence of a morphologically intermediate population in the twilight zone of the cave showing both surface and cave haplotypes suggests the occurrence of gene flow between the surface and cave populations (Schilthuizen et al. 2005). Cases of subterranean taxa from volcanic caves with extant surface relatives living near the cave entrance are also reported from the Galapagos and Canarian archipelagos (see Fig. 3 for a series of Canary Islands representative cave arthropods). Molecular phylogenies have revealed that many of them involve cases of sister-species relationship (Peck & Finston 1993; Arnedo et al. 2007; Contreras-Díaz et al. 2007; Villacorta et al. 2008). A parapatric mode of speciation with ecological shift via resource segregation has been suggested to be applicable in these cases.

In the Tennessee cave salamander species complex of the genus Gyrinophilus (Plethodontidae), Niemiller et al. (2008) used mitochondrial (rrnS and cob) and nuclear (RAG-1) gene genealogies to show that the cave-dwelling populations have recently diverged from the surface species Gyrinophilus porphyriticus via local adaptation and ecological speciation. Cave salamanders have specific traits absent in spring salamanders such as reduced eyes, expanded lateral systems, permanent aquatic life cycle and paedomorphosis, the latter likely promoting premating isolation (Niemiller et al. 2008). Coalescent analyses using the program IM (Hey & Nielsen 2004) found significant evidence of gene flow from the surface G. porphyriticus into the cave G. palleucus necturoides and G. gulolineatus populations, and from the subterranean G. palleucus necturoides to the surface G. porphyriticus, but failed to detect it among cave populations. This suggests a scenario of speciation by divergence with gene flow, a model that has been suggested in other non-subterranean systems (Rice & Hostert 1993; Coyne & Orr 2004).

Speciation within subterranean habitats

Prior to recent molecular genetic studies, it had generally been viewed that subterranean species had reduced phenotypic and genetic diversity because of population bottlenecks resulting from their isolation in cave habitats and adaptation to a stable environment, a hypothesis attributed to Poulson & White (1969). This hypothesis suggested that there would be a limited capacity for subterranean species to undergo adaptive evolution and formation of new species. More recently, however, the occurrence of substantial intraspecific (population level) genetic diversity has been observed within cave species (Stepien et al. 2001; Buhay & Cran-dall 2005; Finlay et al. 2006; Lejeusne & Chevaldonnaé 2006; Guzik et al. 2009). These findings support the alternative view (Barr 1968) that after a reduction in diversity following the initial founder events during
cave colonization, genetic diversity can increase considerably as a subterranean species expands both its range and population size following colonization of new uninhabited subterranean areas.

These earlier studies suggest that sufficient genetic variation can build up within subterranean species to allow for adaptive evolution and speciation. However, to date few studies have examined post-colonization speciation processes within the subterranean habitat. One recent exception consists of a series of studies dealing with sympatric diving beetles from calcrete aquifers in Western Australia (Cooper et al. 2002; Leys et al. 2003; Leys & Watts 2008; Guzik et al. 2009). These calcretes contain at least 12–13 cases of sister-species pairs or triplets of diving-beetle species, where sympatric species fall into non-overlapping size classes. These findings suggest the presence of niche partitioning of beetle species within the calcrete and the potential they evolved by sympatric speciation (Cooper et al. 2002; Leys et al. 2003; Leys & Watts 2008). However, several alternative modes of speciation also may explain this phylogenetic pattern. First, it is possible that sympatric sister species evolved following repeated colonization events from the same ancestral species at different time periods, a process of allopatric divergence. Second, speciation may have occurred within the aquifer by allopatric divergence at a fine spatial scale (referred to as micro-allopatry), possibly through partitioning of the calcrete following water level fluctuations, or by a combination of allopatric/parapatric divergence. In a recent study Guzik et al. (2009) used fine-scale comprehensive sampling of a 3.5 km² grid of bore holes in a single aquifer in Western Australia and examined the mitochondrial gene genealogies of three sympatric sister species of diving beetles. These taxa comprise non-overlapping size classes and form a reciprocally monophyletic clade with respect to other subterranean and surface species of the genus Paroster, suggesting that they evolved within the calcrete from a single common ancestor (Watts & Humphreys 2006; Leys & Watts 2008). The phylogeographic and population analyses undertaken suggested the occurrence of patchy genetic structures (at least in two of the three species) and that demographic histories had involved fluctuations in population size. Isolation by distance in the three species was deduced to have occurred as well. These results, added to the lack of detailed information on niche specialization for the triplet of species, suggest that an allopatric speciation model cannot be ruled out for these species. The dynamics of the environment, however, suggests that micro-allopatry is unlikely to persist for long periods of time and that speciation might have occurred by a combination of micro-allopatric and parapatric divergence (Guzik et al. 2009).

An additional case study suggesting the occurrence of intra-cave speciation involves sympatric troglobiont spider species of the genus Dysdera in the Canary Islands (Fig. 3a). These species often show segregation by body size into groups of large, mid-sized and small species coupled with cheliceral modifications, suggesting specialization to different prey species (Arnedo et al. 2007). The case of the sister-species pair Dysdera hernandezi and Dysdera esquiveli in Tenerife is particularly relevant, as they are very similar both morphologically and genetically (cox1 and rrnL sequences) with the exception of their cheliceral structure. These data suggest that they possibly evolved in sympatry by divergent selection of prey-capture strategies (Arnedo et al. 2007).
The recent use of molecular data has raised the possibility of occurrence of extensive allopatric speciation within the underground habitat, resulting from subterranean radiations (Faille et al. 2010; Ribera et al. 2010). Several beetle groups, Leiodidae (suborder Polyphaga) and Trechinae (suborder Adephaga), appear to represent large monophyletic lineages comprising mostly cave species. Faille et al. (2010) have used a combination of mitochondrial (cox1, cob, rnl1, rnl2, nad1) and nuclear (Small Subunit Ribosomal RNA (SSU), Large Subunit Ribosomal RNA (LSU)) sequences to study the phylogeography and speciation of Trechinae in the Pyrenees. The robust phylogeny obtained for this group demonstrated monophyly for all cave taxa in this mountain range, to the exclusion of all surface and all subterranean species from other geographical areas (Cantabrian and Iberian mountains, Alps). Within the Pyrenees, eastern and western splits were evident, with contrasting evolutionary patterns in each clade apparently related to differences in limestone structure (Fig. 2). In a similar study on Mediterranean Leptodirini (Coleoptera, Leiodidae), Ribera et al. (2010) revealed the presence of several monophyletic cave-dwelling lineages in the mountain ranges of the Iberian Peninsula. These authors proposed an ancient origin for the main Western Mediterranean lineages of Leptodirini, which would have derived from Early-Mid Oligocene (34–25 Ma) ancestors already adapted to subterranean life. Their findings contradict the traditional view that subterranean lineages invariably derive from multiple colonizations by surface ancestors.

Prospects for future research

While phylogeographic and population genetic analyses have greatly contributed to our understanding of the evolutionary and biogeographic history of cave animals, many questions still remain, particularly in the areas of speciation and adaptive and regressive evolution. Recent theoretical developments in speciation modelling (Gavrilets & Vose 2005; Gavrilets et al. 2007; Gravilets & Vose 2007) and massive parallel sequencing techniques (next generation sequencing, Tautz et al. 2010) offer enormous potential in the study of these topics. Similarly, the development of more complex coalescence models (e.g. IMA: Isolation with Migration Models, Hey 2010) using multiple genetic markers to study dispersal and isolation processes will allow a more rigorous assessment of the colonization and population genetic history of subterranean animals.

Gavrilets et al. (2007) have shown that if several factors co-occur, such as few loci controlling local adaptation, intermediate selection and carrying capacity, and if there is a strong effect of non-random mating control loci, parapatric or sympatric (defined as a population where mating is random with respect to the birthplace of the mating partners, Gavrilets 2003, 2004) speciation may be possible. In cave ecosystems, the occurrence of strong selection has been demonstrated for several traits conferring an advantage in the subterranean environment, and these can be rapidly fixed in small populations (see Wilkens 2010 for a review in cave fish). However, recurrent gene flow from surface to cave populations may break co-adapted gene complexes unless there is strong assortative mating, probably via the evolution of behavioural changes in reproductive traits. So-called magic traits (Gavrilets 2004), which are traits under diversifying selection that can simultaneously result in reproductive isolation, provide a basis for speciation to proceed to completion in the presence of gene flow, but their existence in cave animals is not well documented. Paedomorphosis in cave salamanders, where subterranean species retain gills allowing them to live permanently under water, may represent an example of a magic trait, as it also restricts the location of mating, which can occur outside the water in surface salamanders (Niemiller et al. 2008). To develop realistic individual-based models for the study of processes of colonization of subterranean habitats and of subsequent speciation with gene flow in cave animals, good estimates of dispersal, levels of trait selection, number of loci involved, etc. are necessary from an appropriate model system. The latter should comprise cave and surface species preferably displaying parapatric distributions (Gavrilets et al. 2007).

Advances in parallel sequencing technology are starting to change the way we think about many disciplines in molecular biology and evolution (Tautz et al. 2010). Several exciting new approaches using large-scale genome sequencing can be envisaged in subterranean organisms. Some but not all are (i) scanning adaptation and convergence evolution at the genome level, and identifying the genetic basis of these processes, using suitable phylogenetic contrasts of cave and surface species, (ii) the assessment of gene flow across the genome among surface and subterranean taxa to test the adaptive-shift hypothesis, (iii) studying the genetics of phenotypic differences in cave and surface related forms through transcriptional profiling, (iv) the development of tools for rapid biodiversity assessment and detection of cryptic species based on environmental sampling methodologies, and (v) using an across-the-genome multi-gene approach to obtain more reliable molecular dating and estimates of how long it takes for a particular subterranean species to evolve.

For all of these advances to be successful, there is a need for study systems in which both cave and surface populations are of relatively recent origin and where the
key biological, geographical and phylogenetic contexts are well established. Case studies lacking any evidence of the occurrence of allopatric phases would be particularly useful to study the likelihood of ecological speciation through the methodologies mentioned earlier, and this is expected to be more likely to occur among cave troglobionts or stygobionts from the tropics and their relatives.

Acknowledgements

We are very grateful to Bill Humphreys, Brent Emerson and Pedro Oromí for their constructive comments on an early version of the manuscript. The valuable comments and discussions on the paper of three anonymous reviewers and the suggestions of Louis Bertnatchez greatly improved the manuscript. Tom Iliffe, Sergio Montagud, Ignacio Ribera and Pedro Oromí kindly provided photographs for figures. Financial support has been received from the Ministerio de Ciencia e Innovación, Spain (projects CGL2006-01365 to CJ and CGL2009-08256 to DJ, co-financed with EU FEDER funds). CJ visit to the University of Adelaide (Australia) has been funded by the UIB, and to the University of East Anglia (United Kingdom) and Institute of Zoology, Beijing (China) by MEC (PR2009-0231). MTG and SJBC received support for their research from the Australian Research Council (DP0663675 and LP0669062).

References


Carlos Juan research area is molecular evolutionary genetics and phylogeography. Michelle Guzik is an evolutionary biologist with interest on the evolution of invertebrates in isolated aquatic habitats in Australia’s arid region. Damià Jaume works on diversity, comparative morphology and biogeography of copepod and peracarid crustaceans. Steve Cooper’s research area is phylogeography and speciation of Australian fauna, and he has a strong interest in the evolution of subterranean animals from the Australian arid zone.