Author/s:
Frankenberg, SR;Lucas, S;Feigin, CY;Doronina, L;Steffen, R;Hartley, G;Grady, P;Menzies, BR;De Paoli-Iseppi, R;Donnellan, S;Klein, M;Newton, A;Black, JR;Clark, M;Cooper, S;O'Neill, R;Clark, N;Schmitz, J;Pask, AJ

Title:
Unearthing the secrets of Australia's most enigmatic and cryptic mammal, the marsupial mole

Date:
2025-01

Citation:
Frankenberg, S. R., Lucas, S., Feigin, C. Y., Doronina, L., Steffen, R., Hartley, G., Grady, P., Menzies, B. R., De Paoli-Iseppi, R., Donnellan, S., Klein, M., Newton, A., Black, J. R., Clark, M., Cooper, S., O'Neill, R., Clark, N., Schmitz, J. & Pask, A. J. (2025). Unearthing the secrets of Australia's most enigmatic and cryptic mammal, the marsupial mole. Science Advances, 11 (1), https://doi.org/10.1126/sciadv.ado4140.

Persistent Link:
https://hdl.handle.net/11343/356472

GENETICS

Unearthing the secrets of Australia's most enigmatic and cryptic mammal, the marsupial mole

Stephen R. Frankenberg[1]*, Sarah Lucas[2], Charles Y. Feigin[1,3], Liliya Doronina[4,5,6], Raphael Steffen[4], Gabrielle Hartley[7], Patrick Grady[7], Brandon R. Menzies[1], Ricardo De Paoli-Iseppi[8], Stephen Donnellan[9], Mitzi Klein[10], Axel Newton[1], Jay R. Black[11], Michael Clark[8], Steven Cooper[9,12], Rachel O'Neill[7], Nathan Clark[2,13], Jürgen Schmitz[4], Andrew J. Pask[1]*

The marsupial moles are arguably Australia's most enigmatic marsupials. Almost indistinguishable from placental (eutherian) moles, they provide a striking example of convergent evolution. Exploring the genome of the southern marsupial mole, we provide insights into its unusual biology. We show definitively by retrophylogenomic analysis that marsupial moles are most closely related to bandicoots and bilbies (order Peramelemorphia). We find evidence of a marked decline in marsupial mole effective population size, most likely preceding the arrival of humans in regions near its range, and potentially corresponding to periods of climatic change. Our analysis of loss of eye function—an adaptation to subterranean life—reveals a structured order of loss of gene function associated first with the lens, then cone, and finally rod cells. Last, we identify genetic changes suggestive of adaptation to an oxygen-poor environment and of its evolution of partially descended testes.

INTRODUCTION

Marsupial moles (known as *itjaritjari* to the Pitjantjatjara/Yankunytjatjara people) have been well known to the traditional land owners for centuries, where they feature in the *tjukurpa* (creation) stories at Uluru, one of Australia's best known landmarks (*1*). However, when the decomposed specimen of a southern marsupial mole (*Notoryctes typhlops*) was first taxonomically described in 1888 (*2*, *3*), it was considered not to be a marsupial at all, but rather a species basal to the marsupial (Metatheria) and placental (Eutheria) mammals. Later, better quality specimens of *N. typhlops* and its relative, the northern marsupial mole (*Notoryctes caurinus*), confirmed their status as marsupials but they were placed within their own order, Notoryctemorphia, and were assumed to be only distantly related to other marsupials (*4*).

When first formally described, marsupial moles were observed to be completely unlike any other marsupial species, having a unique fossorial nature and bearing a striking resemblance to the eutherian golden mole (Fig. 1). Marsupial moles are small (14 to 18 cm long and weighing 40 to 70 g) and have adaptations to their subterranean environment including eye reduction, modified fore- and hindlimbs, a rostral horny shield, and fused cervical vertebrae to assist in burrowing (*5*, *6*). The presence of two nipples in the backward-facing pouch indicates that only one to two young are produced per litter. Marsupial moles are the only marsupial lacking a scrotum, presumably to minimize drag when moving through sand/earth. The testes undergo partial testicular descent and are located between the abdomen and dermis, cranial to the penis as in other marsupials (*7*). The limited information on their biology has been further confounded by the inability to keep marsupial moles in captivity (*8*).

The northern marsupial mole (*N. caurinus*) is distributed in northwestern Australia and very rarely seen. The southern marsupial mole (*N. typhlops*) is distributed across central and southern Australia, with rare but occasional sightings recorded within the Uluru-Kata Tjuta National Park (*1*). It is unknown whether the northern mole is less abundant than the southern mole or if their less frequent sightings are simply due to the lower abundance of people in their range (*1*). Marsupial moles are known to be preyed on by introduced foxes and cats. Analysis of predator scats in the Tanami Desert showed that 10% of fox scats, 3% of cat scats, and 5% of dingo scats contained mole remnants (*9*). Because of the paucity of information on these animals, the International Union for Conservation of Nature (IUCN 2015) have listed the genus as *least concern*. In the absence of a marsupial mole genome assembly, it has been difficult to determine their current and historical effective population size to support the IUCN classification, or to examine their precise phylogenetic relationship to other marsupials. Furthermore, no information exists on the genetic basis of their unique convergence with eutherian moles despite having over 160 million years of phylogenetic divergence.

Previous efforts to define the marsupial mole's phylogeny have resulted in conflicting evidence for their evolutionary position within Euaustralidelphia—the clade of australidelphian marsupials (*10*). Some morphological analyses supported a close relationship between Notoryctemorphia and Peramelemorphia (*11*). However, this contradicts early DNA hybridization studies (*4*), suggesting that Notoryctemorphia is only distantly related to Peramelemorphia and closer to Diprotodontia. By contrast, sequence-based analyses relying mainly on mitochondrial

[1]School of BioSciences, The University of Melbourne, Parkville, Victoria 3010, Australia. [2]Department of Human Genetics, University of Utah, Salt Lake City, UT, USA. [3]Department of Environment and Genetics, La Trobe University, Bundoora, Victoria 3086, Australia. [4]Institute of Experimental Pathology, ZMBE, University of Münster, 48149 Münster, Germany. [5]Institute for Evolution and Biodiversity, University of Münster, Hüfferstraße 1, Münster 48149, Germany. [6]JICE—Joint Institute for Individualisation in a Changing Environment, University of Münster, Hüfferstraße 1, Münster 48149, Germany. [7]Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT, USA. [8]Department of Anatomy and Physiology, The University of Melbourne, Victoria, Australia. [9]South Australian Museum, North Terrace, Adelaide, South Australia 5000, Australia. [10]Monash Biomedical Imaging, Monash University, Clayton, Victoria, Australia. [11]School of Geography, Earth and Atmospheric Sciences; Melbourne Trace Analysis for Chemical, Earth and Environmental Sciences Platform, The University of Melbourne, Parkville, Victoria 3010, Australia. [12]School of Biological Sciences and Environment Institute, The University of Adelaide, South Australia 5005, Australia. [13]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA, USA.
*Corresponding author. Email: srfr@unimelb.edu.au (S.R.F.); ajpask@unimelb.edu.au (A.J.P.)

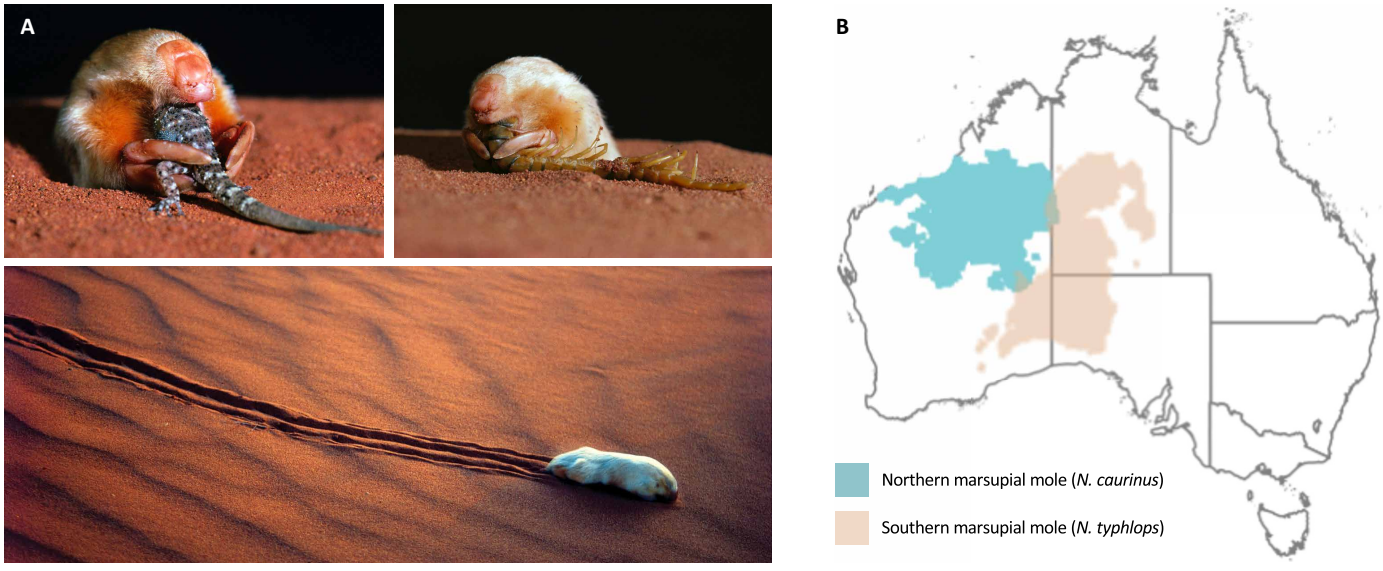

Fig. 1. Marsupial moles and their distribution in central Australia. (**A**) Southern marsupial moles. Image credit: M. Gillam, AUSCAPE. (**B**) Distribution of northern and southern marsupial moles in Australia (distribution data were obtained from the Atlas of Living Australia, accessed 20 June 2024).

sequence data identified *Notoryctes* as a sister group to *Phascogale*, a specific branch within Dasyuromorphia, which also includes the numbat, thylacine, Tasmanian devil, quolls, dunnarts, and antechinuses (*12*). Other sequence-based analyses were inconclusive but tended to merge *Notoryctes* with dasyuromorphs (*13*, *14*) or to place it at the basal position in Agreodontia (a clade including Dasyuromorphia, Peramelemorphia, and Notoryctemorphia) (*15*, *16*). According to a recent genome-scale study of 1550 genomic loci across 18 marsupial families, totaling 867,000 aligned nucleotides (*17*), there is robust support for placing *Notoryctes* as the closest group to Peramelemorphia, consistent with previous morphology studies (*11*, *18*, *19*). Nonetheless, further investigation is needed to establish their exact position within the marsupial family tree.

Marsupial moles resemble several very distantly related eutherian clades of subterranean mammals, including the true moles (family Talapidae), the African golden moles (family Chrysochloridae), the blind mole rats (family Spalacidae), the pocket gophers (Geomyidae), and the naked mole rat (*Heterocephalaus glaber*). All groups share common adaptations to their subterranean niche, such as reduced vision, modified limbs for burrowing, and abdominal testes. The African golden moles are the most analogous to marsupial moles in that they live in arid, sandy habitats and do not form tunnels during burrowing. Instead, both marsupial moles and African golden moles "swim" through the sand, which collapses behind them (*20*). Eutherian moles have evolved unique changes to their hemoglobin as an adaptation to low oxygen in the subterranean environment (*21*).

Here, we report sequencing and assembly of the genome of a female marsupial mole specimen obtained from the South Australia Museum. At 3.85 Gb, it is the largest marsupial genome sequenced to date. By comparing retrotransposon insertions with other marsupial genomes, we confirm the phylogenetic placement of Notoryctemorphia as a sister clade to the Peramelemorphia. Using multiple sequentially Markovian coalescent (MSMC) analysis, we trace the effective population size of southern marsupial moles back 200,000 years, before the arrival of humans in Australia. Comparison of coding sequences of genes involved in eye development identified progressive loss of genes associated first with development of the lens followed by rod and cone cell function. We also describe potential adaptations in genes controlling testicular descent and hemoglobins, which facilitate early development in the hypoxic environment of the pouch.

RESULTS AND DISCUSSION

Genome sequencing and assembly

To produce a chromosome-level genome assembly for the southern marsupial mole (*N. typhlops*), we extracted DNA from a female specimen obtained from the Southern Australian Museum and generated ~50× coverage of long-read sequencing data using Oxford Nanopore Technologies (ONT) R9.1.4 flow cells and Illumina short read data. Contigs were assembled de novo and scaffolded using publicly available Hi-C sequencing data for the southern marsupial mole [National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) accession SRX13785675]. The resulting assembly has a total length of 3.85 Gb, with 10 super-scaffolds and a contig N50 of 356.9 Mb, slightly higher than the *k*-mer–based genome size estimate of 3.56 to 3.61 Gb predicted using Hi-C data. Benchmarking universal single-copy orthologs (BUSCO) analysis identifies 94.2% complete gene models (Table 1). This compares favorably with recent assemblies of other agreodont marsupials such as the eastern quoll and greater bilby (each ~93.5%), though with a relatively high fraction of duplicated orthologs (10.6%) (*22*, *23*).

Using RepeatMasker, we annotated 44.67% of the assembly as repetitive (~1.72 Gb), predominantly composed of long interspersed elements (LINEs) and short interspersed elements (SINEs) at 28.28% and 10.31%, respectively.

Phylogenetic position of the marsupial mole resolved by retrophylogenomics

One of the biggest questions surrounding the marsupial mole is its relationship to other marsupial groups. We took advantage of the

Table 1. Assembly, BUSCO, and Merqury statistics for the marsupial mole genome assembly 1.0.

Assembly statistics	
Number of scaffolds	10
Number of contigs (including unplaced)	693
Total length	3,845,140,653
Largest scaffold	761,411,215
N50	356,873,583
N90	262,881,994
L50	4
L75	8
GC%	37.80
Number of Ns/100 kb	10.64
BUSCO statistics (5.2.2, mammalia_odb10, *n* = 9226)	
Complete	94.2% (8690)
Complete, single copy	83.6% (7715)
Complete, duplicated	10.6% (975)
Fragmented	1.1% (105)
Missing	4.7% (431)
Merqury statistics (version)	
Quality Values (QV) score	36.90
K-mer completeness	93.61%
Per base accuracy rate	99.98%

whole-genome sequence of *N. typhlops* and a virtually homoplasy-free, sequence-independent, retrotransposon presence/absence marker system (*24*) to clarify the enigmatic phylogenetic position of the marsupial mole. The random genomic insertion of a retrotransposon in the common ancestor of species and the subsequent complete fixation of this marker in the population ensure an almost noise-free and phylogenetically correct distribution after speciation. Related species carry an orthologous diagnostic retrotransposon, while less related groups lack this marker in orthologous genomic locations. Hundreds of such diagnostic insertions collected from entire genomes reveal the actual phylogenetic relationships and possible incidences of incomplete lineage sorting (ILS) that appear in cases of rapid consecutive speciations with insufficient intervening time for fixation of markers in the populations.

We extracted 1,071,207, 1,620,648, and 493,756 SINE coordinates from the genomes of the marsupial mole, eastern barred bandicoot, and the Tasmanian devil. The three selected genomes represent the relevant lineages—Notoryctemorphia, Peramelemorphia, and Dasyuromorphia, respectively. We systematically screened for retrotransposon presence/absence patterns among these lineages and eight additional marsupial genomes (see Methods) to resolve the exact phylogenetic place of Notoryctemorphia within Agreodontia. Of these, 436 insertions supported the grouping of Notoryctemorphia plus Peramelemorphia, 25 the grouping of Dasyuromorphia plus Peramelemorphia, and 16 the grouping of Notoryctemorphia plus Dasyuromorphia (Fig. 2 and data S1 and S2). The Kuritzin-Kischka-Schmitz-Churakov (KKSC) insertion significance test (436/25/16 diagnostic markers) significantly supported a shared origin of Notoryctemorphia and Peramelemorphia ($P < 2.1 \times 10^{-156}$). The small number of conflicting signals occurred most probably due to ILS.

We performed additional screenings to verify the position of Notoryctemorphia on a larger scale. The presence of 26 SINE markers in orthologous positions in the Notoryctemorphia, Peramelemorphia, and Dasyuromorphia species and their absence in Diprotodontia confirm significantly (KKSC test, $P < 2.2 \times 10^{-5}$) the monophyly of Agreodontia (Fig. 2; nine conflicting markers for Dasyuromorphia plus Diprotodontia and four conflicting markers for Notoryctemorphia plus Diprotodontia, data S1 and S2).

Previous attempts to identify the position of a marsupial mole using retrotransposons were restricted by the lack of whole-genome data of all agreodontian representatives. They provided only a few signals with conflicts (*10*, *25*). Here, we performed the first whole-genome retrotransposon screening of representatives of all three lineages of Agreodontia, applying a similar screening strategy for diagnostic SINE insertions that recently testified the reunion of Australasian possums (*26*). Our data support the position of the marsupial mole within a monophyletic clade—Agreodontia. The common ancestor of Agreodontia persisted for less than 2 million years before further diversifications (*17*). ILS often accompanies such rapid radiations and most probably explains the few conflicting SINE signals that suggested placement of Diprotodontia within Agreodontia.

The 436 SINE loci shared by Notoryctemorphia and Peramelemorphia provide highly significant evidence for their close grouping, as opposed to only 25 markers shared by Dasyuromorphia and Peramelemorphia, and only 16 markers supporting a Notoryctemorphia-Dasyuromorphia sister group relationship. Our genome-wide, retrotransposon presence/absence results agree with earlier morphological data (*11*, *19*, *27*) and the previous

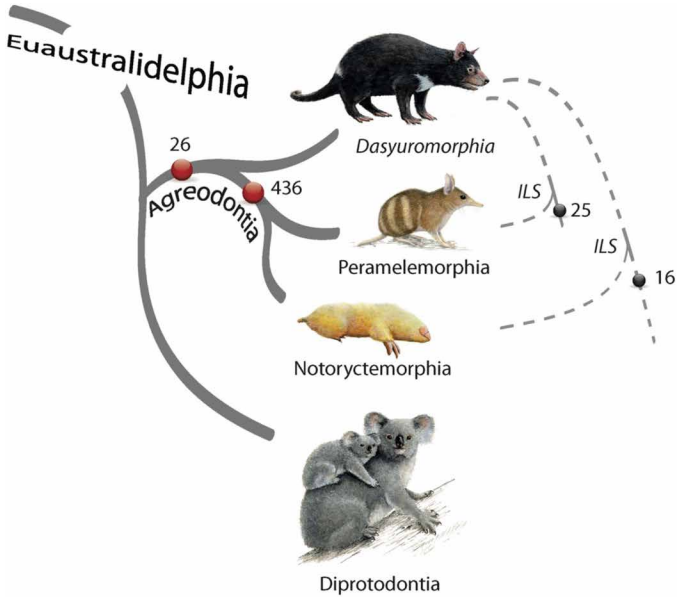

Fig. 2. Retrophylogenomic position of the marsupial mole among Euaustralidelphia. The orthologous positions of 26 diagnostic insertions verify a shared origin of Agreodontia comprising Dasyuromorphia, Peramelemorphia, and Notoryctemorphia (sequence information of Peramelemorphia was available only for seven loci, see Methods and data S2; conflicting agreodontian markers are not shown). Four hundred thirty-six orthologous diagnostic insertions provide significant evidence of Notoryctemorphia as the sister group to Peramelemoprhia. The underlying genome-level dataset also revealed a few conflicting signals of 25 insertions that suggest a grouping of Dasyuromorphia plus Peramelemorphia and 16 of grouping Notoryctemorphia plus Dasyuromorphia (dashed lines), indicating incidences of ILS. Image credit: J. Baldur Hlidberg, fauna.is.

sequence-based analysis of 1550 genomic loci (*17*). However, they do not support the conclusions derived from other previous sequence-based analyses (*12, 15, 16*), possibly because these data were based on too few sequences or faulty signals from mitochondrial sequences. Our data have finally resolved the position of the marsupial mole and, based on the sequence analyses of Duchêne *et al.* (*17*), the common ancestor of marsupial moles and bandicoots existed around 60 million years ago (Ma). The procedure for extracting nearly conflict-free diagnostic retrotransposon markers involves the automated removal of low-quality loci during *n*-way processing, resulting in a robust phylogenetic signal that is independent of genome completeness. The likelihood of hundreds of diagnostic retrotransposon insertion loci arising independently from parallel insertions into identical genomic loci in mammals is negligible (*24*, *28*). Phylogenetically diagnostic retrotransposon loci, together with genome-scale sequence data from 1,550 genomic exons (*17*), support the close relationship between the marsupial mole and Peramelemorphia. These two approaches offer independent significance by providing virtually homoplasy-free, almost selectively neutral genome-scale insertion markers and selectively relevant genome-scale exon sequences.

Demographic history of the marsupial mole suggests declining numbers

Almost nothing is known about the demographics of this marsupial, which is seldom seen and almost impossible to detect in population surveys. Moreover, little is known about how this species may have been affected by major climatic and biogeographical shifts during Australia's recent history. Using MSMC analyses with MSMC2 (*29*, *30*), we reconstructed the effective population size (*Ne*) for the southern marsupial mole during the late Pleiostocene and early Holocene.

Because of the cryptic nature of this species, no data exist on its generation time or on germline mutation rate, which are key parameters for scaling *Ne* and time in years in historical demographic inferences. However, recently published mutation rates for the related Tasmanian devil (*Sarcophilus harrisii*) and the South American opossum (*Monodelphis domestica*) fall within a relatively narrow range compared to those of eutherian mammals (5.95×10^{-9} and 4.60×10^{-9}, respectively), suggesting that these may reflect a reasonable estimate for the marsupial mole (*31*). To explore the potential impact that alternative parameter choices may have on the inferred demographic trajectory, we scaled the results of our analyses using each of these mutation rates, together with a plausible range of generation times (from 1 to 3 years) based on those of other extant Australian marsupials (Fig. 3 and fig. S1) (*32*). Here, we focus chiefly on the trajectory inferred using the mutation rate of the related Tasmanian devil and an intermediate value of generation time (2 years) (Fig. 3).

Under these parameters, our analysis indicates that before ~100 thousand years ago (ka), effective population size was fairly stable but appear to have dropped sharply just over 70 ka, corresponding roughly with the beginning of marine isotope stage 4 (MIS4; Fig. 3). MIS4 (approximately 71 ka to 59 ka) was a period of cooler temperatures and dropping sea levels, marking the beginning of the Last Glacial Period (LGP) (*33*). It is therefore possible that climatic changes during this period may have influenced marsupial mole

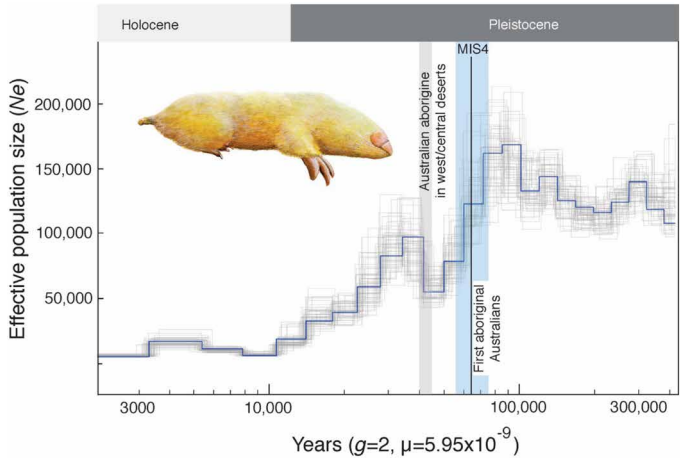

Fig. 3. Demographic history of the marsupial mole. This figure shows the demographic history of *N. typhlops* during the late Pleiostocene and early Holocene, as a step plot and scaled using a generation time of 2 years ($g = 2$) and a mutation rate (μ) of 5.95×10^{-9}. Demographic inference from the whole genome is shown in blue, while gray lines represent 50 bootstrap replicates generated by randomly resampling the data. Major biogeographical events during this time frame are shown for context. Marine isotope stage 4 (MIS4; ~71 to 59 ka) is shown as a vertical blue bar, initial human arrival in Australia (~65 ka) is marked with a black vertical line, and the timeframe in which humans began inhabiting areas adjacent to the marsupial mole's range (~45 to 40 ka) is indicated with a vertical gray bar. Image credit: J. Baldur Hlidberg, fauna.is.

populations, although there is substantial variation in the correspondence between MIS4 and marsupial mole decline depending on parameter choices (fig. S1). Interestingly though, across the parameter range explored, declines in marsupial mole *Ne* precede the local arrival of humans in areas surrounding its range [roughly 45 to 40 ka (*34*)] and, under some scenarios, their arrival on the continent (fig. S1) (*35*). While the climatic conditions of MIS4 are themselves proposed as factors influencing human migration patterns, our analyses do not indicate that human arrival itself is responsible for declining *Ne* in the marsupial mole. Nevertheless, variation in estimates under different parameter sets, as well as the use of a single genome assembly, leave room for further exploration of marsupial mole demography (*36*).

Analysis of eye loss in marsupial moles and the progressive evolution of vision loss in mammals

Despite radiating into a diverse array of environments, mammalian vision and related proteins have remained remarkably conserved (*37*, *38*). However, there are several independent instances where mammals no longer benefit strongly from vision, such as living in a subterranean vision (*39*–*43*). These species have since experienced progressive loss of vision, diminished eye structures, and an increased number of changes in both retina and lens in protein-coding genes and in noncoding regions (*41*, *44*–*55*). While the evolution of vision loss has been described in several placental mammal species, it has received less attention in marsupials.

The ideal marsupial to confirm the translatability of these previous findings is the subterranean marsupial mole. Morphologically, marsupial moles have no external evidence of the eye (*56*). Internally, the eye is highly degenerate with no evidence of a lens and loss of the optic nerve. A pigment layer remains and is suggested to be a remnant of the retina; however, rod and cone cells are not present in this tissue (*56*). We examined the marsupial mole genome to determine the degree of degradation in protein-coding genes related to vision (see Methods and data S3) and to compare the order in which constraint was lost between different eye regions. While previous analyses in placental mammals describe degradation in multiple retinal proteins (*57*) and in a single lens protein separately (*58*, *59*), we compared the rate of change between retinal cell types and lens fiber cells, allowing us to draw conclusions about when specific ocular functions across the whole eye were released from constraint. Knowledge of that timing provides clues as to the steps that occurred during their transition to subterranean life.

To determine if ocular genes are under relaxed constraint, we relied on the expectation that, upon release, nonsynonymous and synonymous substitutions should occur at equal frequencies, as quantified using their d_N and d_S rates (*60*, *61*). As such, the d_N/d_S ratio is expected to increase until it approaches 1, the ratio of neutrality; however, this increase occurs gradually over a branch, perhaps over millions of years, due to past constraint. As anticipated, the general upward shift of d_N/d_S ratios in ocular genes shows that they are under reduced selection, relative to other marsupials, and the shift is specific to ocular genes when compared to a random set of 650 non-ocular genes (one-tailed *t* test: lens $P = 4.5 \times 10^{-9}$, cone $P = 0.037$, rod $P = 0.00064$) (data S3 and fig. S2). This suggests that the orthologous genes responsible for placental mammal vision play similar roles in marsupials. While there is high uncertainty in the point estimate of d_N/d_S for a given gene, treating the genes as a sample within a cell type allows us to make some statements about when constraint was relaxed.

We next estimated how long ago each of the vision-related genes lost constraint using a method used in the field that contrasts d_N/d_S between the marsupial mole and other marsupials (Fig. 4, B to D, and data S3) (*62*–*64*). This method makes multiple assumptions that should limit estimates to the lower bounds of time since constraint was released. It assumes that constraint was released completely and instantaneously rather than a gradual release. Therefore, it is a lower bound because any deviation from those assumptions would return an earlier time point.

Despite these limitations, this estimate, even as a lower bound, is useful to place in the context of evolutionary events. When applied to these data, we calculated that the lens genes were released from constraint at least 16 Ma, cone photoreceptors at least 9 Ma, and rod

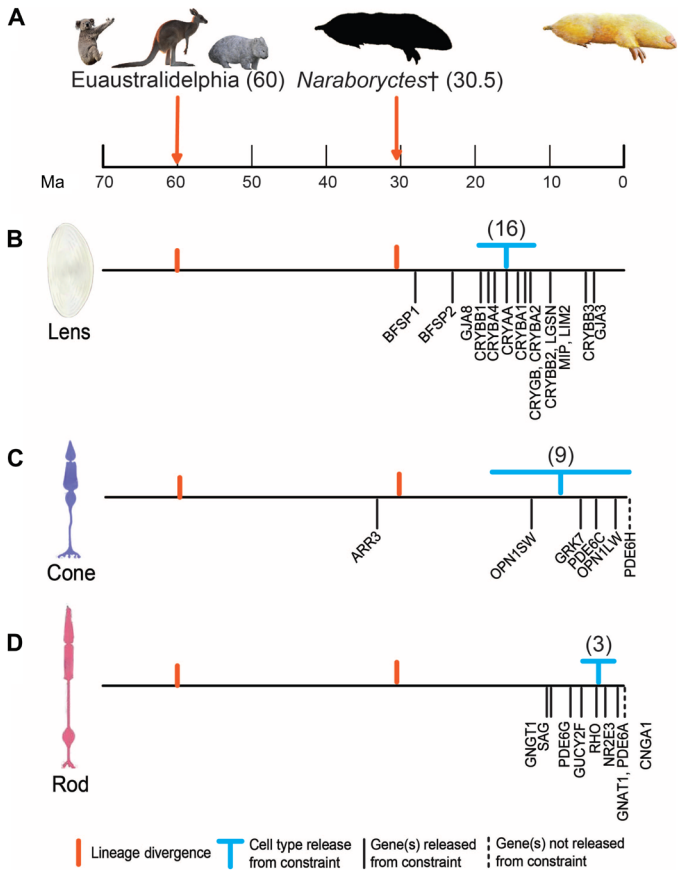

Fig. 4. Progression of loss of constraint among vision-related genes. The images in this figure represent timelines of (**A**) the marsupial mole lineage or (**B** to **D**) tissue-specific events over the last 70 million years with two notable events from (A) shared among all the timelines. (B to D) Solid or dashed lines below the timelines indicate the best estimated time at which the gene lost constraint. The "T" symbol above the timelines indicates the lower-bound estimate of when constraint of that tissue was lost. The tissue loss of constraint was determined based on the weighted average of each individual gene's predicted release from constraint weighted by the length of that gene and is indicated above the T symbol between parentheses. The horizontal bar of the T symbol represents the weighted 95% confidence interval. Image credits: koala image, M. Predavec (CC0 1.0 Universal Deed); kangaroo image, N. Strickland (CC0 1.0 Universal Deed); wombat image, W. Koch (CC0 1.0 Universal Deed); marsupial mole image, J. Baldur Hlidberg, fauna.is. Lens, cone, and rod images were released into the public domain as a licensed U.S. Government Work by the National Eye Institute.

photoreceptors at least 3 Ma (Fig. 4, B to D). These timings are large-ly in agreement with a recent study by Springer *et al.* (*62*), which examined conservation of genes critical for vision in the marsupial mole (Table 2). While there were 22 genes that both groups studied, the Springer *et al.* (*62*) study examined an additional 15 genes and did not screen for pleiotropy, while our study reviewed an addition-al 8 genes and filtered out those with expression in non- ocular tissues. These differences allowed Springer *et al.* (*62*) to investigate how different processes evolved within the marsupial mole in contrast to our study, which examined changes that occurred within specific cell types. We also note that our interpretation of the dates from both studies as a lower bound makes sense, because they would fall after speciation of the marsupial mole and *Naraboryctes*, an extinct relative whose fossil evidence suggests was also subterranean (Fig. 4A).

The timing estimates are also useful to study the relative ordering of loss of constraint in the different cell types. They strongly support that the lens was released from constraint before the rods (weighted *t* test, $P = 1.93 \times 10^{-5}$) with the relative timing of the cones being placed between them, although remaining inconclusive (cones versus lens, $P = 0.24$; cones versus rods, $P = 0.28$). These findings suggest that a loss of constraint on lens genes could be an early indicator of vision loss in subterranean species, and we propose that release of constraint on lens fiber proteins could be used to identify other lineages beginning the process of visual degradation. The relatively late relaxation of constraint on the rod photoreceptors suggests that they remain or remained important for limited vision; since rods are important for vision in low light, their longer retention fits with a progressive model of visual degradation, as the species perhaps gradually became more subterranean.

Mutations in *RXFP2* may explain why the marsupial mole became the only marsupial with partially descended testes

The marsupial mole shows only limited testicular descent and is the only marsupial without a scrotum. We verified the subcutaneous location of the testes using micro–computed tomography (CT) and magnetic resonance imaging (MRI) scanning (Fig. 5A and fig. S3). We then examined the nucleotide sequence of testicular descent genes including *insulin–like factor 3* (*INSL3*) and its receptor, *relaxin/insulin–like family peptide 2* (*RXFP2*). These two genes are essential for the trans-abdominal phase of testicular descent in therian mammals and have been shown to be lost or mutated in afrotherian mammals that are secondarily testicond (have lost testicular descent), such as the lesser hedgehog tenrec, Cape golden mole, Cape elephant shrew, and manatee (*65*–*67*).

While the marsupial mole appears to have a functional *INSL3* gene with some nonsynonymous substitutions relative to other mammals, we have identified potentially disabling mutations in the *RXFP2* gene including a premature stop codon in exon 16 (Fig. 5B and fig. S4), and 55 additional substitutions located mostly downstream of the premature stop codon, of which 27 produce nonsynonymous amino acid changes relative to other marsupials, while 12 produce amino acid changes at sites highly conserved across all mammals. The premature stop codon in exon 16 truncates the amino acid sequence by 165 residues, which corresponds to a large portion of the transmembrane domain and C terminus that stimulate cytoplasmic signal transduction and adenosine 3′,5′-monophosphate (cAMP) production (*68*). Thus, the predicted marsupial mole RXFP2 protein is likely completely nonfunctional.

We compared the ratio of nonsynonymous to synonymous (d_N/d_S) mutations for evidence of relaxed selection in both *INSL3* and *RXFP2* using the RELAX and PAML programs. While both results were not significant for relaxed selection, the d_N/d_S ratio for *RXFP2* in the *Notoryctes* branch was higher (0.19) than the average (0.14) for a range of mammals including platypus, indicating more relaxed selection at this locus in marsupial moles. Similar to degradation of the eye loss genes described above, our data suggest the loss of RXFP2 function may have resulted in partial testicular descent and loss of the scrotum in the marsupial mole.

Duplication of ω-hemoglobin may facilitate life in the pouch for a subterranean marsupial

Omega hemoglobin (ω-hemoglobin) is expressed in the developing marsupial embryo just before birth and in early neonates, suggesting that it helps with the oxygen demands for the newborn in the low oxygen environment of the pouch (*69*). This low oxygen environment is further confounded in the marsupial mole, which already inhabits a low oxygen environment, being a subterranean species. Many subterranean placental mammals show novel hemoglobin adaptations to facilitate their survival in hypoxic conditions (*70*, *71*).

We examined the hemoglobin gene cluster in the marsupial mole and found a duplication of the ω-hemoglobin gene (Fig. 6). This is significant, given that it is the only known duplication of ω-hemoglobin described in any mammal. We suggest that the additive pressure of a subterranean lifestyle and high oxygen demands for young in the pouch may have led to the duplication of this gene to support early development.

The cryptic lifestyle of marsupial moles has greatly limited the extent to which fundamental aspects of their biology have been

Table 2. Comparison of findings between this study and the Springer *et al.* (*62*) study. Dates shown are when each cell group was released from constraint in millions of years ago (Ma).

	Time released from constraint (Ma)		
Cell type	**This study**	**Table 4 of Springer *et al.* (*62*)**	
Lens	16	17.84	Lens/cornea development
Cone	9	38.85	Cone phototransduction recovery
		5.38	Cone phototransduction activation
Rod	3	3.39	Rod phototransduction activation
		1.52	Rod phototransduction recovery

studied. By sequencing, assembling, and analyzing the genome of a marsupial mole species, we have unearthed many unique features of this unusual mammal. In accordance with genome-scale sequence data (*27*), we have resolved their phylogeny, with conclusive evidence that marsupial moles are most closely related to the Perameloformorphia, followed by Dasyuromorphia. We provide evidence that the estimated effective population size remained stable until roughly 100,000 years ago but has since experienced a significant decline to less than a thousand individuals recently. This conclusion is based on our analysis using parameters drawn from the Tasmanian devil as a reference point. By comparing the rate of relaxation of constraint over time among subsets of eye-specific genes, our analysis reveals a structured order of loss of functional eye components during the evolutionary transition to a fully subterranean life. Finally, we identified genetic adaptations to enable survival in the oxygen-poor environment and mutations leading to testicondy, another common feature of subterranean life. Our study provides fundamental insights into the gain and loss of specific traits in mammals to facilitate life underground. Furthermore, this work has immensely expanded the knowledge base for the marsupial mole and provides a new framework for the conservation and management of this unique and fascinating species.

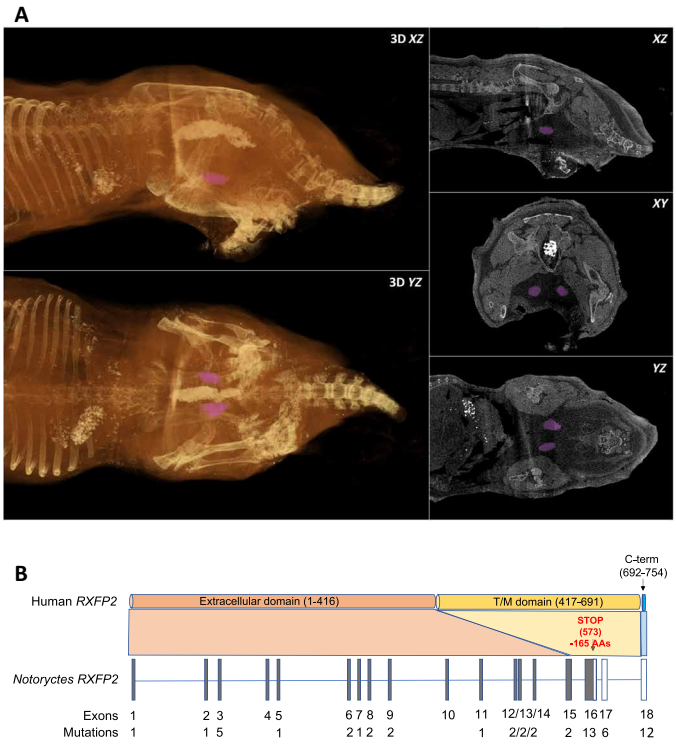

Fig. 5. Mutation in the testicular descent gene *RXFP2* in the marsupial mole. (**A**) MRI scan of a male southern marsupial mole. Left panels represent 3D renderings, and the right panels are *X*, *Y*, and *Z* slices. Testes are highlighted in magenta. (**B**) *RXFP2* of the marsupial mole appears to be under somewhat relaxed selection given a significant number of nonsynonymous mutations, including a premature stop codon in exon 16 that truncates the encoded protein by 165 amino acids. Many of the mutations are present at or downstream of the stop codon (exons 16 to 18) in a region that may not interfere with continued gene function but may be in the process of being lost, given the lack of a scrotum in the marsupial mole and only partial testicular descent. White bars in exons denote mutations.

METHODS

Sequencing

Tissues from a single individual captured on 8 January 2015, in the Waite Creek area of the Northern Territory, Australia, were stored frozen in the Australian Biological Tissue Collection (ABTC voucher 07570), South Australian Museum. Muscle tissue was then used to extract genomic DNA using the Nanobind Tissue Big DNA Kit (Circulomics). Two separate batches of sequencing were performed using PromethION flow cells (ONT). The first was performed by the Australian Genome Resource Facility on a library constructed with the Ligation Sequencing Kit (SQK-LSK109, ONT). Additional libraries constructed with the Ligation Sequencing Kit (SQK-LSK109, ONT) or the ONT/Circulomics Ultra-long kits (SQK-ULK001 and NB-900-601-01, respectively) were sequenced at the Walter and Eliza Hall Institute sequencing facility. FASTQ files were generated from the FAST5 data based on a super-accuracy model using the Guppy basecaller software v6.0.7 (ONT, UK).

Illumina short-read sequencing was performed by the Ramaciotti Centre for Genomics (UNSW Sydney, Australia) on a NovaSeq 6000 platform. All raw data have been made available through NCBI SRA (BioProject PRJNA1193092).

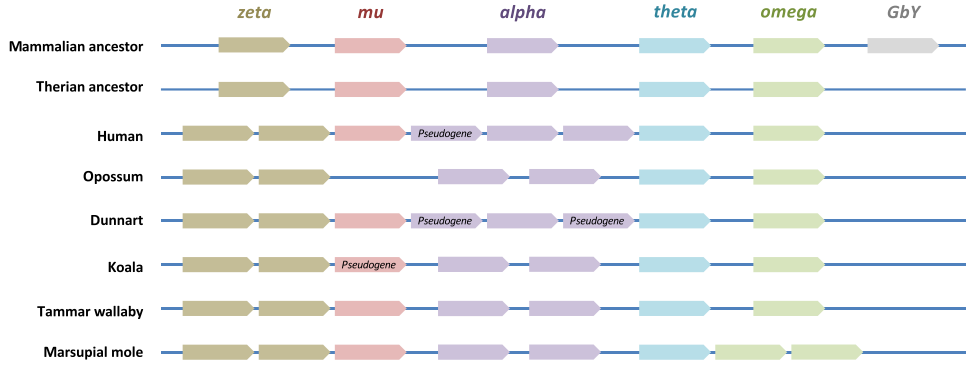

Fig. 6. Comparison of hemoglobin gene synteny within the α-hemoglobin cluster among selected mammals. Ancestral mammalian and therian gene compositions are after Patel *et al.* (*106*). The marsupial mole is the only species examined with two copies of the omega gene.

Genome assembly

To remove excess mitochondrial sequences, ONT reads were first mapped against the published marsupial mole mitochondrial genome assembly (NC_006522.1) using minimap2 (v 2.22-r1101) (*72*) with default parameters. Aligned reads were subsequently removed using samtools. Contigs were assembled from the remaining reads using Flye (v2.9) (*73*) with parameter –nano-hq. An initial round of polishing using ONT reads was performed using medaka consensus (v1.5.0) (*74*) with the r941_prom_sup_g507 base calling model set. Next, Illumina short reads were aligned against the contigs using bwa-mem2 (v2.2.1) (*75*) with parameter -M. Read mapping in proper pairs (samtools -f 3) was sorted and provided as input to pilon (v1.2.4) (*76*) for one round of short read polishing with parameters --diploid --chunksize 10000000 --fix snps,indels,gaps and --minqual 15. Unpurged haplotypes were then removed using purge_dups (*77*). Briefly, ONT reads were realigned to the contigs using minimap2 with parameter -x map-ont, and contigs were separately self-aligned using minimap2 with parameter -x asm5 and -DP. The included pbcstat and calcuts utilities used to identify coverage cutoffs were then applied to filter contigs using purge_dups, and a purged fasta file was generated with the get_seqs utility.

Juicer (1.6) (*78*) and 3D-DNA (180922) were used to process publicly available Hi-C reads for the southern marsupial mole (NCBI SRA accession SRX13785675) (https://www.ncbi.nlm.nih.gov/sra/SRX13785675) and scaffold the draft assembly using default parameters. Juicebox with Assembly Tools (1.11.08) was used to manually review the produced scaffolds and correct errors. BBMap was used to filter contigs less than 3 kb in length, and 149.11 Gb of Oxford Nanopore reads was used to gap-fill the scaffolded assembly with TGS-GapCloser (1.0.1).

QUAST (5.2.0) was used to assess genome statistics, and BUSCO (5.2.2) was used to estimate gene completeness using the mammalia_odb10 database. Repeat annotations were performed with RepeatMasker (4.1.3) using a combined Dfam 3.6 and Repbase (20181026) database. Subsequently, RepeatMasker was used with a custom repeat library to identify KERV long terminal repeat elements and other marsupial-derived satellites not present in the repeat databases. The resulting repeat annotations were combined and summarized using the RepeatMasker utility buildSummary.pl.

Meryl and Merqury were used to assess the base-level accuracy of the assembly (*79*). To do so, the raw nanopore reads were input into Meryl to create a *k*-mer database with a *k*-mer size of 22 to assess the presence or absence of *k*-mers in the assembly. A filter was added at this step to remove any *k*-mers with abundance of 1, since these are more likely to be nanopore errors (A. Rhie, personal communication). Merqury generated the quality score for the assembly using this *k*-mer database, as shown in Table 1.

To assess structurally error-prone regions in the assembly, bedtools genomecov (*80*) was used to calculate areas in which nanopore coverage did not support the assembly contiguity. The nanopore reads were mapped using Winnowmap, a program built upon Minimap2 but requiring unique *k*-mer anchors to improve mapping quality (*72*, *81*). Approximately 9035 blocks in the assembly were noted to have zero coverage, representing a minimum total of 200 Mb, suggesting that they are potentially misassembled regions. These blocks were intersected with the genes included in relevant analyses, and no overlap was found between them.

Using the publicly available Hi-C reads described above, GenomeScope (v2.0) (*82*) was used to estimate the expected marsupial mole genome size based on *k*-mer frequencies with a *k*-mer size of 21 and a maximum *k*-mer coverage of 10,000,000.

Defining the phylogeny of the marsupial mole

We soft-masked repeats from the genomes of the southern marsupial mole (*N. typhlops*, Notoryctemorphia), eastern barred bandicoot (*Perameles gunnii*, Peramelemorphia; Oz Mammals Genomics, https://data.bioplatforms.com/omg-10x-processed-illumina/bpa-omg-10x-processed-illumina-102_100_100_40136-hmcktbbxx), the Tasmanian devil (*S. harrisii*, Dasyuromorphia; mSarHar1.11, NCBI, https://www.ncbi.nlm.nih.gov/), and koala (*Phascolarctos cinereus*, phaCin_unsw_v4.1, NCBI) using the Repeat Masker (replacing repeat sequences from a standard "Metatheria" repeat library with small letters; https://www.repeatmasker.org/RepeatMasker/). The repeat masked genomes were used to generate the following pairwise genome alignments [multiple 2-ways (*83*)]: (i) southern marsupial mole/eastern barred bandicoot, (ii) southern marsupial mole/Tasmanian devil, (iii) southern marsupial mole/koala, (iv) eastern barred bandicoot/southern marsupial mole, (vi) eastern barred bandicoot/Tasmanian devil, 6) eastern barred bandicoot/koala, (vii) Tasmanian devil/southern marsupial mole, (viii) Tasmanian devil/eastern barred bandicoot, (ix) Tasmanian devil/koala, (x) koala/southern marsupial mole, and (xi) koala/Tasmanian devil. These 2-ways were integrated into *n*-way (*83*) to analyze the presence/absence patterns of diagnostic SINE elements (WSINE1, WSINE1a, WALLSI1, WALLSI1a, WALLSI3 MAR1a_Mdo, and MAR1b_Mdo) that were active across euaustralidelphian species during the expected diversification of the marsupial mole about 60 Ma (*17*, *84*). To identify the phylogenetic position of the marsupial mole within Agreodontia, we used the SINE coordinates of the southern marsupial mole, eastern barred bandicoot, and Tasmanian devil genomes reciprocally as references for detecting and determining the presence or absence in the remaining species and the outgroups [koala (*P. cinereus*), wallaby (*Notamacropus eugenii*, me-1 k.fasta, DNA Zoo, https://www.dnazoo.org/), and monito del monte (*Dromiciops gliroides*, mDroGli1.pri, NCBI)]. Analyses were performed via *n*-way (*83*) and local blast (https://www.ncbi.nlm.nih.gov/guide/howto/run-blast-local/). We investigated all possible position of the marsupial mole within Agreodontia: (i) southern marsupial mole (+), eastern barred bandicoot (+), Tasmanian devil (−); (ii) southern marsupial mole (+), Tasmanian devil (+), eastern barred bandicoot (−); (iii) eastern barred bandicoot (+), Tasmanian devil (+), southern marsupial mole (−). Furthermore, we verified the monophyly of Agreodontia by performing additional *n*-way screenings and using the marsupial mole, Tasmanian devil, and koala as reference genomes. We detected the following signals: (i) southern marsupial mole (+) Tasmanian devil (+) koala (−); (ii) southern marsupial mole (+) koala (+) Tasmanian devil (−); and (iii) Tasmanian devil (+) koala (+), southern marsupial mole (−), taking monito del monte as an outgroup. Considering the highly significant sister group relationships of Notoryctemorphia and Peramelemorphia (see Results), we assumed them as one monophyletic lineage and excluded Peramelemorphia from the *n*-way screening. If possible, we supplemented all alignments with local blast information of the eastern barred bandicoot. We accepted loci only where Notoryctemorphia and Peramelemorphia showed the same signal or missing sequence information for the latter. This increased the critical number of informative loci for the rapidly diversified Agreodontia clades.

For all *n*-way screenings, we extracted and analyzed alignments of potentially informative SINEs plus their 1-kb flanking regions for

manual verification of SINE orthology in related species. For Notoryctemorphia and Peramelemorphia, there was only one genome per group available. However, we supplemented alignments with additional available dasyuromorphian species {yellow-footed antechinus (*Antechinus flavipes*, AdamAnt, NCBI), fat-tailed dunnart (*Sminthopsis crassicaudata,* Sminthopsis_crassicaudata_HiC.fasta, DNA Zoo), and thylacine [*Thylacinus cynocephalus* (*85*)]}. For the Agreodontia screenings, we supplemented alignments with additional diprotodontian species: Leadbeater's possum (*Gymnobelideus leadbeateri*; LBP_v1, NCBI) and ground cuscus (*Phalanger gymnotis*; pg-2 k.fasta, DNA Zoo). Unique SINE loci that were consistently present or absent in the representatives of each group, absent in the orthologous positions in outgroups, and shifted less than 3 nucleotides were considered phylogenetic diagnostic markers. SINE loci with conflicting signals within dasyuromorphia, missing presence/absence of Diprodontia, or lacking the monito del monte outgroup were removed. Furthermore, because of some inconsistency of thylacine signals within Dasyuromorphia, we removed all loci undefined in the Tasmanian tiger. A statistical evaluation was performed using the KKSC insertion significance test (*86*) (http://retrogenomics.uni-muenster.de:3838/KKSC_significance_test/).

Reconstructing the demographic history of the marsupial mole by MSMC2

We mapped the *Notoryctes* sequence reads to our genome assembly (EBB_10x-rails.scaffolds.fa) using the Burrows-Wheeler Alignment algorithm bwa mem (default parameters, version 0.7.17), which efficiently links short sequencing reads to large reference genomes (*87*). We then computed the average read depth of the resulting BAM file via the SAMtools command samtools depth with the setting -a so that it outputs all positions [Li and Durbin (*87*), version 1.10]. The resulting coverages per base were summarized and divided by the overall number of positions, resulting in an average coverage of 50.7832X. We excluded all scaffolds of putative X chromosome origin using a gene annotation overlap procedure. Briefly, gene annotations from the current version of the Tasmanian devil genome (GCF_902635505.1/mSarHar1.11) were first lifted over to the marsupial mole assembly using liftoff v1.6.1 (parameter: -d 4). The percentage of genes originating from each Tasmanian devil chromosome were then tabulated across each marsupial mole scaffold. Marsupial mole scaffolds for which 75% or more genes originated from the Tasmanian devil X chromosome (NC_045432.1) were conservatively excluded from demographic analyses. We also filtered out all scaffolds smaller than 50 kb to avoid including low-quality regions in our overall assessment.

These filtering steps left 554 putative autosomal scaffolds. Single-nucleotide polymorphism (SNP) variants were called using the SAMtools commands bcftools mpileup (-q 20 -Q 20 -C 50) and bcftools call (-c -V indels). The mpileup program generates genotype likelihoods for every genomic location with coverage measurement, while the calling program bcftools does the actual calling. The -q and -Q arguments set the minimum limits for mapping the qualities of alignments and bases, respectively; those calls falling short of these exclusion limits were skipped by mpileup. Furthermore, -C adjusts the mapping quality for reads containing excessive mismatches, with 50 as the recommended value. For the call command, we decided to use -V to skip indels and to call only SNPs, while -c forces the command to use the original bcftools calling method instead of the available alternative due to the latter being more suited for rare-variant calling. Additionally, we used the bamcaller.py script, provided by MSMC-tools, as part of the pipeline to generate the necessary bed files (mask.bed.gz) for each scaffold directly from the mpileup standard input data. The only argument it requires is the previously calculated mean average coverage.

To generate the MSMC input (multihetsep) files for all scaffold VFC and mask.bed.gz files, we used the generate_multihetsep.py script of MSMC-tools with default settings. Low complexity and repetitive regions were masked with an additional BED file previously generated using the RED repeat masking software. We ran MSMC2 with 100 iterations and default settings for the final calculation. The resulting file was then plotted via a custom R script. Additionally, we used the multihetsep_bootstrap.py script, which is also part of MSMC-tools, with default settings to generate 50 bootstrap datasets consisting of 30 pseudochromosomes each created by chopping the multihetsep file into blocks of 5 kb and randomly sampling them. The MSMC2 bootstrap results of these datasets were then plotted alongside our main result. We used the recently published parent-offspring trio mutation rate of the Tasmanian devil ($\mu = 5.95 \times 10^{-9}$) (*31*) and a generation time of $g = 2$ years for plotting due to the phylogenetic proximity of both species, despite lacking specific data on mutation rate or generation time for *Notoryctes*. Recognizing that these are rough estimates for the parameters, we also present additional plots using $g = 1$ and $g = 3$ as lower and upper bounds for the generation time and $\mu = 4.6 \times 10^{-9}$, the mutation rate of *M. domestica*, as an alternative rate for a less closely related species. This approach provides an overview of the potential variability in interpreting the MSMC data (see fig. S1). Increasing generation times move data points further into the past, whereas increasing mutation rates reduce the timeframe and decrease the effective population size values.

Analysis of vision genes

Using the assembled genome, we generated protein models using Braker2 (v 2.1.6) (*88*), genome threader (v 1.7.1) (*89*), and Liftoff (*90*). Braker2 was run to detect proteins of any evolutionary distance (--softmasked --epmode). All vertebrate proteins (v100, orthodatabase v 10) were downloaded, extracted, and concatenated into a single fasta file for ProtHint to generate protein hints. Genome threader ran using the softmasked version of the genome, using the protein models from Tasmanian Devil v 1.11 (GCF_902635505.1) and with options -gff3out and -skipalignmentout. The output from all the programs was concatenated together (see data S4).

To identify what subset of proteins has eye expression, all 9782 proteins detected by liquid chromatography–tandem mass spectrometry (LC-MS/MS) across 784 human samples were used (*91*). These proteins were identified using UniProt IDs. Of those proteins, 9729 were converted to HGNC names by UniProt mapping tool for comparison with Human Protein Atlas data (*92*). To identify proteins that are only expressed in the eye, those proteins were subjected to additional screening to remove proteins with multi-tissue expression. Proteins that were characterized using immunocytochemistry/immunofluorescence by Human Protein Atlas (The Human Protein Atlas version 21.1 and Ensembl version 103.38) to have any level of protein expression in any tissue, except in the eye (anterior section of eye), retina, or testis with low to medium expression, were removed from consideration (*93*). For proteins that only had eye (anterior section) expression, this alone was sufficient for inclusion as there is typically no RNA expression done in the lens or lens

fiber cells. As we were unable to filter based on RNA expression, some genes included in our lens gene set could be pleiotropic and therefore led to an underestimation of when the lens fiber cells were released from constraint. Two additional proteins (LIM2 and LENEP) were also included as they are known lens proteins through literature review (*94*, *95*).

For proteins expressed in the retina, additional criteria were needed for the protein's inclusion. When comparing tissue RNA levels, the protein could not have more than 10% of the retina RNA expression in any other tissue. In addition when single cell RNA data were available, the protein could not have more than 10% of the maximum RNA expression level in retinal cell type (bipolar, cone photoreceptor, horizontal, Muller glia, or rod photoreceptor cell) in any other cell type, except for early and later spermatid, in which case the RNA expression had to be less than 50% of the maximum RNA expression of a retinal cell type (*96*). Higher RNA expression in spermatids was allowed as it is shown to have a permissive transcriptome (*97*). Some retinal proteins are expressed in both photoreceptor cell types. Proteins were considered rod/cone photoreceptor proteins if their single cell expression was at least 5.5× greater than the other photoreceptor cell type.

To increase the representation of marsupial mammals, we identified several publicly available marsupial genomes with gene models, protein models, and annotation files along with the marsupial mole and dunnart ones generated in this study (data S3). As a first pass, we used inParanoid (5.0, default parameters) to identify orthologous proteins between human and the different marsupial species (*98*, *99*). Protein models were used instead of nucleotide models because inParanoid had more success in identifying orthologs based on their protein sequence. This resulted in 124 missing genes (out of a possible 720). To recover some of these missing genes, TBLASTN (B LAST v2.11.0) was used to identify orthologous regions (*100*, *101*). If a nucleotide model was present in this region, it was manually examined and nucleotides that appeared to cause a frameshift were removed. If there was no nucleotide model, then a nucleotide model was assembled based on exons identified from TBLASTN. These manually created nucleotide models were translated and added to the protein model list. As a second and final pass, inParanoid was run again the same as the first time with one notable exception. As no marsupial orthologs for *CRX* were recovered by inParanoid using the human ortholog, they were replaced with the gray short-tailed opossum (*M. domestica*) orthologs. This recovered all but 32 genes.

Orthologous codon sequences were determined using the protein results from InParanoid and identifying the corresponding gene model through the annotation file. The codon models of the identified protein models were collated into fasta files based on orthologous group manually to ensure that only the coding sequence portion was included. Codon sequences were aligned using PRANK (170427, codon option) (*102*). They were visually inspected using Aliview (*103*). If less than 50% of the nucleotides were homologous between one species' codon model and the rest of the orthologous codon models, then another codon model with the same inParanoid score was used. If there was not another codon model, it was removed from the alignment. With the alignments, we estimated branch-specific d_N/d_S or omega (ω) for the marsupial mole using the codeml program within PAML (v 4.9, seqtype = 1, CodonFreq = 3, model = 2, getSE = 1, cleandata = 0) (*17*). To calculate time since pseudogenization, we used equation 3 from Meredith *et al.* (*63*) reproduced here

$$T_p = T - \frac{T(\omega_m - \omega_p)}{\omega_f - \omega_p}$$

where T_p is the time since pseudogenization; T is the time since the lineages diverged from another, which is predicted to be 60 Ma; ω_m is the d_N/d_s calculated for the marsupial mole branch; ω_p is the predicted d_N/d_s on the pseudogenic/marsupial mole branch and assumed to be 1; and ω_f is the calculated d_N/d_s for the rest of the marsupial tree (*63, 104*). One key assumption of this equation is that all constraint has been lost on that lineage. If not all constraint was lost, then the release from constraint could have happened earlier.

In tissues with at least five genes, tissue loss of constraint was determined using a weighted average of when each gene is predicted to have lost constraint, weighted by the gene's length. Weighted averages were used so that smaller genes would not have an oversized weight in determining the whole tissue's timing. The weighted 95% confidence intervals were calculated using the wtd.var function in the Hmisc package (v5.1-1).

$$\sqrt{\frac{\text{wtd. var}(x, \text{weights})}{n}} \times 1.96$$

where x is the time each gene was released from constraint, weights is the length of that gene, and n is the number of genes in that tissue.

To determine if the tissue was released from constraint, a random set of 650 genes was chosen to form a null set. A marsupial mole ortholog had to be present in all those genes and genes that had a predominantly ocular function were excluded (data S3). These alignments were formed as described above and run through PAML using the same parameters as noted above except getSE = 0. For each set of genes (null, lens, cone, and rod), the difference between the d_N/d_S of the marsupial mole from the background species was calculated. A one-tailed t test compared if average d_N/d_S difference of the null set was significantly less than the average d_N/d_S difference for any of the three individual tissue sets (*105*).

Analyses of testicular descent genes

Mutations observed in the *Notoryctes RXFP2* genome sequence were confirmed by individual polymerase chain reactions (PCRs) using exon-specific primers (table S1), gel purification, and Sanger sequencing. PCRs included a 2× SYBR Green Master Mix (Promega, USA), nuclease-free water, 10 μM of forward and reverse primer, and 1 μl of genomic DNA in 20-μl total volume. PCR conditions included 95°C at 5 min followed by 35 cycles of 95°C for 15 s, a specific annealing temperature (table S1), 72°C for 30 s, then 72°C for 5 min. Samples were run on a 1% agarose gel and bands of correct size, excised under ultraviolet (UV) light, and purified using a gel PCR purification kit (Qiagen, Germany). Samples were sequenced at the Australian Genome Research Facility using an internal sequencing primer.

Micro-CT scanning

Micro-CT scanning was performed with a phoenix nanotom m (Waygate Technologies) operated using xs control and the phoenix

datos|x acquisition software. The marsupial mole specimen was wrapped in bubble wrap and mounted in a polyvinyl chloride pipe for scanning. An x-ray energy of 90 kV and 440 μA was used with a 0.5-mm Al filter and molybdenum target. Three 10-min scans were conducted collecting 1199 projections through a full 360-degree rotation of the specimen at three vertical positions to capture the full length of the specimen at a resolution of 25 μm. The first two scans were collected with the specimen's snout downward, and then the specimen was inverted to collected the tail scan. Volume reconstruction of the micro-CT data was performed using the phoenix datos|x reconstruction software applying an inline median filter and region of interest filter during reconstruction and data output in a 32-bit format. Data were imported to Avizo (Thermo Fisher Scientific) to register and merge the three scan volumes together to form a single stitched volume reconstruction of the specimen. The final volume was converted to a 16-bit format for further segmentation and visualization. Reconstructed tiff stacks were loaded into DragonFly (Object Research Systems; Montreal, Quebec, Canada) for visualization, volume rendering, and segmentation under a noncommercial license.

MRI scanning

MRI scanning was performed on a male marsupial mole held at the Tieg's Museum in the School of BioSciences, University of Melbourne. MRI experiments were performed on a Bruker BioSpec 9.4 T scanner (Bruker BioSpin GmbH, Ettlingen, Germany) with a Bruker 86-mm volume coil as the transmitter and Bruker 4-array rat heart coil as the receiver to acquire axial, sagittal, and coronal images using two-dimensional (2D) FLASH sequence. Imaging parameters were as follows: repetition time = 2000 ms, echo time = 3 ms, matrix = 256 × 256, field of view = 51.2 × 51.2 mm^2, slice thickness = 0.5 mm, slice numbers = 80, flip angle = 40°, averages = 3.

Analyses of hemoglobin genes

Hemoglobin coding sequences were identified by performing nucleotide BLAST searches of the marsupial mole assembly using as queries the coding sequences from the nearest related species available for each orthologue (*zeta*, *mu*, *alpha*, *theta*, and *omega*). Orthology of marsupial mole genes was confirmed by reciprocal BLAST searches.

Supplementary Materials

The PDF file includes:
Supplementary Text
Figs. S1 to S4
Table S1
Legends for data S1 to S4

Other Supplementary Material for this manuscript includes the following:
Data S1 to S4

REFERENCES AND NOTES

1. K. Bennison, J. Clayton, R. Godfree, C. Pavey, M. Wilson, Surfacing behaviour and ecology of the marsupial mole (*Notoryctes typhlops*) at Uluru-Kata Tjuta National Park. *Aust. Mammal.* **36**, 184–188 (2014).
2. E. C. Stirling, Preliminary notes on a new Australian mammal. *Trans. R. Soc. S. Aust.* **1888**, 21–24 (1888).
3. E. C. Stirling, Description of a new genus and species of marsupialia, Notoryctes typhlops. *Trans. R. Soc. S. Aust.* **14**, 154–187 (1891).
4. J. A. W. Kirsch, M. S. Springer, F.-J. Lapointe, DNA-hybridisation studies of marsupials and their implications for metatherian classification. *Aust. J. Zool.* **45**, 211–280 (1997).
5. J. Benshemesh, K. Johnson, "Biology and conservation of marsupial moles (*Notoryctes*)" in *Predators with Pouches: The Biology of Carnivorous Marsupials* (CSIRO Publishing, 2003), p. 505.
6. J. Benshemesh, "Itjaritjari" in *The Mammals of Australia* (Reed New Holland, 2008).
7. E. C. Stirling, Communications to the Society. *Proc. R. Soc. Lond.* , 327–329 (1891).
8. D. Howe, Observations on a captive marsupial mole, Notoryctes typhlops. *Aust. Mammal.* **1**, 361–365 (1975).
9. R. Paltridge, Occurrence of the marsupial mole (Notoryctes typhlops) remains in the faecal pellets of cats, foxes and dingoes in the Tanami Desert, N.T. *Aust. Mammal.* **20**, 427–429 (1998).
10. M. A. Nilsson, G. Churakov, M. Sommer, N. V. Tran, A. Zemann, J. Brosius, J. Schmitz, Tracking marsupial evolution using archaic genomic retroposon insertions. *PLOS Biol.* **8**, e1000436 (2010).
11. I. Horovitz, M. R. Sánchez-Villagra, A morphological analysis of marsupial mammal higher-level phylogenetic relationships. *Cladistics* **19**, 181–212 (2003).
12. M. S. Springer, M. Westerman, J. R. Kavanagh, A. Burk, M. O. Woodburne, D. J. Kao, C. Krajewski, The origin of the Australasian marsupial fauna and the phylogenetic affinities of the enigmatic monito del monte and marsupial mole. *Proc. Biol. Sci.* **265**, 2381–2386 (1998).
13. J. D. Retief, C. Krajewski, M. Westerman, R. J. Winkfein, G. H. Dixon, Molecular phylogeny and evolution of marsupial protamine P1 genes. *Proc. Biol. Sci.* **259**, 7–14 (1995).
14. C. Krajewski, L. Buckley, M. Westerman, DNA phylogeny of the marsupial wolf resolved. *Proc. Biol. Sci.* **264**, 911–917 (1997).
15. M. J. Phillips, P. A. McLenachan, C. Down, G. C. Gibb, D. Penny, Combined mitochondrial and nuclear DNA sequences resolve the interrelations of the major Australasian marsupial radiations. *Syst. Biol.* **55**, 122–137 (2006).
16. K. J. Mitchell, R. C. Pratt, L. N. Watson, G. C. Gibb, B. Llamas, M. Kasper, J. Edson, B. Hopwood, D. Male, K. N. Armstrong, M. Meyer, M. Hofreiter, J. Austin, S. C. Donnellan, M. S. Y. Lee, M. J. Phillips, A. Cooper, Molecular phylogeny, biogeography, and habitat preference evolution of marsupials. *Mol. Biol. Evol.* **31**, 2322–2330 (2014).
17. D. A. Duchêne, J. G. Bragg, S. Duchêne, L. E. Neaves, S. Potter, C. Moritz, R. N. Johnson, S. Y. W. Ho, M. D. B. Eldridge, Analysis of phylogenomic tree space resolves relationships among marsupial families. *Syst. Biol.* **67**, 400–412 (2018).
18. J. A. W. Kirsch, Szalay, F. S. 1994. Evolutionary History of the Marsupials and an Analysis of Osteological Characters, Cambridge University Press, New York, xii + 481 pp. ISBN 0–521–44169-2. Price (hardbound), $84.95. *J. Mammal.* **76**, 1283–1288 (1995).
19. R. M. D. Beck, N. M. Warburton, M. Archer, S. J. Hand, K. P. Aplin, Going underground: Postcranial morphology of the early miocene marsupial mole naraboryctes philcreaseri and the evolution of fossoriality in notoryctemorphians. *Mem. Mus. Vic.* **74**, 151–171 (2016).
20. K. A. Johnson, "The mammals of Australia" in *The Mammals of Australia*, R. Strahan, Ed. (New Holland Publishers Sydney, 1998), pp. 409–4011.
21. K. L. Campbell, M. J. Gaudry, K. He, H. Suzuki, Y. P. Zhang, X. L. Jiang, R. E. Weber, Altered hemoglobin co-factor sensitivity does not underlie the evolution of derived fossorial specializations in the family Talpidae. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **224**, 150–155 (2018).
22. G. A. Hartley, S. R. Frankenberg, N. M. Robinson, A. J. MacDonald, R. K. Hamede, C. P. Burridge, M. E. Jones, T. Faulkner, H. Shute, K. Rose, R. Brewster, R. J. O'Neill, M. B. Renfree, A. J. Pask, C. Y. Feigin, Genome of the endangered eastern quoll (Dasyurus viverrinus) reveals signatures of historical decline and pelage color evolution. *Commun. Biol.* **7**, 636 (2024).
23. C. J. Hogg, R. J. Edwards, K. A. Farquharson, L. W. Silver, P. Brandies, E. Peel, M. Escalona, F. R. Jaya, R. Thavornkanlapachai, K. Batley, T. M. Bradford, J. K. Chang, Z. Chen, N. Deshpande, M. Dziminski, K. M. Ewart, O. W. Griffith, L. Marin Gual, K. L. Moon, K. J. Travouillon, P. Waters, C. M. Whittington, M. R. Wilkins, K. M. Helgen, N. Lo, S. Y. W. Ho, A. Ruiz Herrera, R. Paltridge, J. A. Marshall Graves, M. Renfree, B. Shapiro, K. Ottewell, R. Kiwirrkurra, K. Belov, Extant and extinct bilby genomes combined with Indigenous knowledge improve conservation of a unique Australian marsupial. *Nat. Ecol. Evol.* **8**, 1311–1326 (2024).
24. L. Doronina, O. Reising, H. Clawson, D. A. Ray, J. Schmitz, True homoplasy of retrotransposon insertions in primates. *Syst. Biol.* **68**, 482–493 (2019).
25. S. Gallus, A. Janke, V. Kumar, M. A. Nilsson, Disentangling the relationship of the Australian marsupial orders using retrotransposon and evolutionary network analyses. *Genome Biol. Evol.* **7**, 985–992 (2015).
26. L. Doronina, C. Y. Feigin, J. Schmitz, Reunion of Australasian possums by shared SINE insertions. *Syst. Biol.* **71**, 1045–1053 (2022).
27. F. S. Szalay, *Evolutionary History of the Marsupials and an Analysis of Osteological Characters* (Cambridge Univ. Press, 2009).

28. L. Doronina, L. Ogoniak, J. Schmitz, Homoplasy of retrotransposon insertions in toothed whales. *Genes* **14**, 1830 (2023).
29. S. Schiffels, R. Durbin, Inferring human population size and separation history from multiple genome sequences. *Nat. Genet.* **46**, 919–925 (2014).
30. S. Schiffels, K. Wang, MSMC and MSMC2: The multiple sequentially markovian coalescent. *Methods Mol. Biol.* **2090**, 147–166 (2020).
31. L. A. Bergeron, S. Besenbacher, J. Zheng, P. Li, M. F. Bertelsen, B. Quintard, J. I. Hoffman, Z. Li, J. St Leger, C. Shao, J. Stiller, M. T. P. Gilbert, M. H. Schierup, G. Zhang, Evolution of the germline mutation rate across vertebrates. *Nature* **615**, 285–291 (2023).
32. A. Burbidge, P. Harrison, J. Woinarski, *The Action Plan for Australian Mammals 2012* (CSIRO Publishing, 2014), p. 1053.
33. P. De Deckker, L. J. Arnold, S. van der Kaars, G. Bayon, J. B. W. Stuut, K. Perner, R. L. dos Santos, R. Uemura, M. Demuro, Marine Isotope Stage 4 in Australasia: A full glacial culminating 65,000 years ago - Global connections and implications for human dispersal. *Quat. Sci. Rev.* **204**, 187–207 (2019).
34. R. Tobler, A. Rohrlach, J. Soubrier, P. Bover, B. Llamas, J. Tuke, N. Bean, A. Abdullah-Highfold, S. Agius, A. O'Donoghue, I. O'Loughlin, P. Sutton, F. Zilio, K. Walshe, A. N. Williams, C. S. M. Turney, M. Williams, S. M. Richards, R. J. Mitchell, E. Kowal, J. R. Stephen, L. Williams, W. Haak, A. Cooper, Aboriginal mitogenomes reveal 50,000 years of regionalism in Australia. *Nature* **544**, 180–184 (2017).
35. C. Clarkson, Z. Jacobs, B. Marwick, R. Fullagar, L. Wallis, M. Smith, R. G. Roberts, E. Hayes, K. Lowe, X. Carah, S. A. Florin, J. McNeil, D. Cox, L. J. Arnold, Q. Hua, J. Huntley, H. E. A. Brand, T. Manne, A. Fairbairn, J. Shulmeister, L. Lyle, M. Salinas, M. Page, K. Connell, G. Park, K. Norman, T. Murphy, C. Pardoe, Human occupation of northern Australia by 65,000 years ago. *Nature* **547**, 306–310 (2017).
36. J. Balme, S. O'Connor, S. Fallon, New dates on dingo bones from Madura Cave provide oldest firm evidence for arrival of the species in Australia. *Sci. Rep.* **8**, 9933 (2018).
37. G. H. Jacobs, Evolution of colour vision in mammals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 2957–2967 (2009).
38. P. Vopalensky, Z. Kozmik, Eye evolution: Common use and independent recruitment of genetic components. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 2819–2832 (2009).
39. C. A. Emerling, M. S. Springer, Genomic evidence for rod monochromacy in sloths and armadillos suggests early subterranean history for Xenarthra. *Proc. R. Soc. Lond. B Biol. Sci.* **282**, 20142192 (2015).
40. R. W. Meredith, J. Gatesy, C. A. Emerling, V. M. York, M. S. Springer, Rod monochromacy and the coevolution of cetacean retinal opsins. *PLOS Genet.* **9**, e1003432 (2013).
41. R. Partha, B. K. Chauhan, Z. Ferreira, J. D. Robinson, K. Lathrop, K. K. Nischal, M. Chikina, N. L. Clark, Subterranean mammals show convergent regression in ocular genes and enhancers, along with adaptation to tunneling. *eLife* **6**, e25884 (2017).
42. A. Sadier, K. T. J. Davies, L. R. Yohe, K. Yun, P. Donat, B. P. Hedrick, E. R. Dumont, L. M. Dávalos, S. J. Rossiter, K. E. Sears, Multifactorial processes underlie parallel opsin loss in neotropical bats. *eLife* **7**, e37412 (2018).
43. H. Zhao, S. J. Rossiter, E. C. Teeling, C. Li, J. A. Cotton, S. Zhang, The evolution of color vision in nocturnal mammals. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 8980–8985 (2009).
44. F. D. Carmona, R. Jiménez, J. M. Collinson, The molecular basis of defective lens development in the Iberian mole. *BMC Biol.* **6**, 44 (2008).
45. R. Cernuda-Cernuda, J. M. García-Fernández, M. C. M. Gordijn, P. H. M. Bovee-Geurts, W. J. DeGrip, The eye of the african mole-rat Cryptomys anselli: To see or not to see? *Eur. J. Neurosci.* **17**, 709–720 (2003).
46. G. Esquiva, A. Avivi, J. Hannibal, Non-image forming light detection by melanopsin, Rhodopsin, and long-middlewave (L/W) cone opsin in the subterranean blind mole Rat, Spalax ehrenbergi: Immunohistochemical characterization, Distribution, and connectivity. *Front. Neuroanat.* **10**, 61 (2016).
47. M. Glösmann, M. Steiner, L. Peichl, P. K. Ahnelt, Cone photoreceptors and potential UV vision in a subterranean insectivore, the European mole. *J. Vis.* **8**, 23 (2008).
48. N. V. Nikitina, B. Maughan-Brown, M. J. O'Riain, S. H. Kidson, Postnatal development of the eye in the naked mole rat (Heterocephalus glaber). *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* **277A**, 317–337 (2004).
49. L. Peichl, A. E. Chavez, A. Ocampo, W. Mena, F. Bozinovic, A. G. Palacios, Eye and vision in the subterranean rodent cururo (Spalacopus cyanus, octodontidae). *J. Comp. Neurol.* **486**, 197–208 (2005).
50. L. Peichl, P. Němec, H. Burda, Unusual cone and rod properties in subterranean African mole-rats (Rodentia, Bathyergidae). *Eur. J. Neurosci.* **19**, 1545–1558 (2004).
51. Y. Quax-Jeuken, S. Bruisten, H. Bloemendal, W. W. De Jong, Evolution of crystallins: Expression of lens-specific proteins in the blind mammals mole (Talpa europaea) and mole rat (Spalax ehrenbergi). *Mol. Biol. Evol.* **2**, 279–288 (1985).
52. C. E. Schleich, A. Vielma, M. Glösmann, A. G. Palacios, L. Peichl, Retinal photoreceptors of two subterranean tuco-tuco species (Rodentia, Ctenomys): Morphology, topography, and spectral sensitivity. *J. Comp. Neurol.* **518**, 4001–4015 (2010).
53. M. Jiang, L. Shi, X. Li, Q. Dong, H. Sun, Y. Du, Y. Zhang, T. Shao, H. Cheng, W. Chen, Z. Wang, Genome-wide adaptive evolution to underground stresses in subterranean mammals: Hypoxia adaption, immunity promotion, and sensory specialization. *Ecol. Evol.* **10**, 7377–7388 (2020).
54. J. G. Roscito, K. Sameith, G. Parra, B. E. Langer, A. Petzold, C. Moebius, M. Bickle, M. T. Rodrigues, M. Hiller, Phenotype loss is associated with widespread divergence of the gene regulatory landscape in evolution. *Nat. Commun.* **9**, 4737 (2018).
55. J. G. Roscito, K. Subramanian, R. Naumann, M. Sarov, A. Shevchenko, A. Bogdanova, T. Kurth, L. Foerster, M. Kreysing, M. Hiller, Recapitulating evolutionary divergence in a single cis-regulatory element is sufficient to cause expression changes of the lens gene Tdrd7. *Mol. Biol. Evol.* **38**, 380–392 (2021).
56. G. Sweet, Contribution to our knowledge of the anatomy of Notoryctes typhlops, Stirling. Part III. The eye. *Q. J. Microsc. Sci.* **50**, 547–572 (1906).
57. C. A. Emerling, M. S. Springer, Eyes underground: Regression of visual protein networks in subterranean mammals. *Mol. Phylogenet. Evol.* **78**, 260–270 (2014).
58. A. Avivi, A. Joel, E. Nevo, The lens protein α-B-crystallin of the blind subterranean mole-rat: High homology with sighted mammals. *Gene* **264**, 45–49 (2001).
59. W. Hendriks, J. Leunissen, E. Nevo, H. Bloemendal, W. W. de Jong, The lens protein alpha A-crystallin of the blind mole rat, Spalax ehrenbergi: Evolutionary change and functional constraints. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 5320–5324 (1987).
60. M. Kimura, Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature* **267**, 275–276 (1977).
61. W. H. Li, *Molecular Evolution* (Sinauer Associates, 2006).
62. M. S. Springer, C. A. Emerling, J. Gatesy, Three blind moles: Molecular evolutionary insights on the tempo and mode of convergent eye degeneration in *Notoryctes typhlops* (southern marsupial mole) and two chrysochlorids (golden moles). *Genes (Basel)* **14**, (2023).
63. R. W. Meredith, J. Gatesy, W. J. Murphy, O. A. Ryder, M. S. Springer, Molecular decay of the tooth gene Enamelin (ENAM) mirrors the loss of enamel in the fossil record of placental mammals. *PLOS Genet.* **5**, e1000634 (2009).
64. K. Nagasawa, T. Kitano, Pseudogenization of the hair-related genes PADI3 and S100A3 in cetaceans and hippopotamus amphibius. *J. Mol. Evol.* **91**, 745–760 (2023).
65. J. I. Park, J. Semyonov, C. L. Chang, W. Yi, W. Warren, S. Y. Hsu, Origin of INSL3-mediated testicular descent in therian mammals. *Genome Res.* **18**, 974–985 (2008).
66. V. Sharma, T. Lehmann, H. Stuckas, L. Funke, M. Hiller, Loss of RXFP2 and INSL3 genes in Afrotheria shows that testicular descent is the ancestral condition in placental mammals. *PLOS Biol.* **16**, e2005293 (2018).
67. S. Zimmermann, G. Steding, J. M. Emmen, A. O. Brinkmann, K. Nayernia, A. F. Holstein, W. Engel, I. M. Adham, Targeted disruption of the Insl3 gene causes bilateral cryptorchidism. *Mol. Endocrinol.* **13**, 681–691 (1999).
68. R. A. Bathgate, R. Ivell, B. M. Sanborn, O. D. Sherwood, R. J. Summers, International union of pharmacology LVII: Recommendations for the nomenclature of receptors for relaxin family peptides. *Pharmacol. Rev.* **58**, 7–31 (2006).
69. D. Wheeler, R. Hope, S. B. Cooper, G. Dolman, G. C. Webb, C. D. Bottema, A. A. Gooley, M. Goodman, R. A. Holland, An orphaned mammalian beta-globin gene of ancient evolutionary origin. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1101–1106 (2001).
70. F. Li, Z. Qiao, Q. Duan, E. Nevo, Adaptation of mammals to hypoxia. *Animal Model Exp. Med.* **4**, 311–318 (2021).
71. M. Li, D. Pan, H. Sun, L. Zhang, H. Cheng, T. Shao, Z. Wang, The hypoxia adaptation of small mammals to plateau and underground burrow conditions. *Animal Model Exp. Med.* **4**, 319–328 (2021).
72. H. Li, Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
73. M. Kolmogorov, J. Yuan, Y. Lin, P. A. Pevzner, Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
74. https://github.com/nanoporetech/medaka.
75. M. Vasimuddin, S. Misra, H. Li, S. Aluru, in *2019 IEEE International Parallel and Distributed Processing Symposium* (IPDPS). (2019), pp. 314-324.
76. B. J. Walker, T. Abeel, T. Shea, M. Priest, A. Abouelliel, S. Sakthikumar, C. A. Cuomo, Q. Zeng, J. Wortman, S. K. Young, A. M. Earl, Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLOS ONE* **9**, e112963 (2014).
77. D. Guan, S. A. McCarthy, J. Wood, K. Howe, Y. Wang, R. Durbin, Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics* **36**, 2896–2898 (2020).
78. N. C. Durand, M. S. Shamim, I. Machol, S. S. Rao, M. H. Huntley, E. S. Lander, E. L. Aiden, Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
79. A. Rhie, B. P. Walenz, S. Koren, A. M. Phillippy, Merqury: Reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol.* **21**, 245 (2020).
80. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

81. C. Jain, A. Rhie, H. Zhang, C. Chu, B. P. Walenz, S. Koren, A. M. Phillippy, Weighted minimizer sampling improves long read mapping. *Bioinformatics* **36**, i111–i118 (2020).
82. T. R. Ranallo-Benavidez, K. S. Jaron, M. C. Schatz, GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **11**, 1432 (2020).
83. G. Churakov, F. J. Zhang, N. Grundmann, W. Makalowski, A. Noll, L. Doronina, J. Schmitz, The multicomparative 2-n-way genome suite. *Genome Res.* **30**, 1508–1516 (2020).
84. A. Zemann, G. Churakov, S. Donnellan, F. Grützner, F. Zhao, J. Brosius, J. Schmitz, Ancestry of the Australian termitivorous numbat. *Mol. Biol. Evol.* **30**, 1041–1045 (2013).
85. C. Y. Feigin, A. H. Newton, L. Doronina, J. Schmitz, C. A. Hipsley, K. J. Mitchell, G. Gower, B. Llamas, J. Soubrier, T. N. Heider, B. R. Menzies, A. Cooper, R. J. O'Neill, A. J. Pask, Genome of the Tasmanian tiger provides insights into the evolution and demography of an extinct marsupial carnivore. *Nat. Ecol. Evol.* **2**, 182–192 (2018).
86. A. Kuritzin, T. Kischka, J. Schmitz, G. Churakov, Incomplete lineage sorting and hybridization statistics for large-scale retroposon insertion data. *PLOS Comput. Biol.* **12**, e1004812 (2016).
87. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
88. T. Brůna, K. J. Hoff, A. Lomsadze, M. Stanke, M. Borodovsky, BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genom. Bioinform.* **3**, lqaa108 (2021).
89. G. Gremme, V. Brendel, M. E. Sparks, S. Kurtz, Engineering a software tool for gene structure prediction in higher organisms. *Inf. Softw. Technol.* **47**, 965–978 (2005).
90. A. Shumate, S. L. Salzberg, Liftoff: Accurate mapping of gene annotations. *Bioinformatics* **37**, 1639–1643 (2021).
91. M. T. Ahmad, P. Zhang, C. Dufresne, L. Ferrucci, R. D. Semba, The human eye proteome project: Updates on an emerging proteome. *Proteomics* **18**, –e1700394 (2018).
92. H. Huang, P. B. McGarvey, B. E. Suzek, R. Mazumder, J. Zhang, Y. Chen, C. H. Wu, A comprehensive protein-centric ID mapping service for molecular data integration. *Bioinformatics* **27**, 1191–1190 (2011).
93. M. Uhlén, L. Fagerberg, B. M. Hallström, C. Lindskog, P. Oksvold, A. Mardinoglu, Å. Sivertsson, C. Kampf, E. Sjöstedt, A. Asplund, I. M. Olsson, K. Edlund, E. Lundberg, S. Navani, C. A. K. Szigyarto, J. Odeberg, D. Djureinovic, J. O. Takanen, S. Hober, T. Alm, P. H. Edqvist, H. Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J. M. Schwenk, M. Hamsten, K. Von Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. Von Heijne, J. Nielsen, F. Pontén, Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
94. G. J. Maher, G. C. Black, F. D. Manson, Focus on molecules: Lens intrinsic membrane protein (LIM2/MP20). *Exp. Eye Res.* **103**, 115–116 (2012).
95. Y. Wen, G. Sachs, C. Athmann, A novel lens epithelium gene, LEP503, is highly conserved in different vertebrate species and is developmentally regulated in postnatal rat lens. *Exp. Eye Res.* **70**, 159–168 (2000).
96. M. Karlsson, C. Zhang, L. Méar, W. Zhong, A. Digre, B. Katona, E. Sjöstedt, L. Butler, J. Odeberg, P. Dusart, F. Edfors, P. Oksvold, K. von Feilitzen, M. Zwahlen, M. Arif, O. Altay, X. Li, M. Ozcan, A. Mardonoglu, L. Fagerberg, J. Mulder, Y. Luo, F. Ponten, M. Uhlén, C. Lindskog, A single–cell type transcriptomics map of human tissues. *Sci. Adv.* **7**, (2021).
97. M. Soumillon, A. Necsulea, M. Weier, D. Brawand, X. Zhang, H. Gu, P. Barthès, M. Kokkinaki, S. Nef, A. Gnirke, M. Dym, B. DeMassy, T. S. Mikkelsen, H. Kaessmann, Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Rep.* **3**, 2179–2190 (2013).
98. E. Persson, E. L. L. Sonnhammer, InParanoid-DIAMOND: Faster orthology analysis with the InParanoid algorithm. *Bioinformatics* **38**, 2918–2919 (2022).
99. M. Remm, C. E. V. Storm, E. L. L. Sonnhammer, Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.* **314**, 1041–1052 (2001).
100. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
101. E. M. Gerts, Y. K. Yu, R. Agarwala, A. A. Schäffer, S. F. Altschul, Composition-based statistics and translated nucleotide searches: Improving the TBLASTN module of BLAST. *BMC Biol.* **4**, 41 (2006).
102. A. Löytynoja, Phylogeny-aware alignment with PRANK. *Methods Mol. Biol.* **1079**, 155–170 (2014).
103. A. Larsson, AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **30**, 3278–3276 (2014).
104. R. W. Meredith, M. Westerman, M. S. Springer, A phylogeny of Diprotodontia (Marsupialia) based on sequences for five nuclear genes. *Mol. Phylogenet. Evol.* **51**, 554–571 (2009).
105. U. Knief, W. Forstmeier, Violating the normality assumption may be the lesser of two evils. *Behav. Res. Methods* **53**, 2576–2590 (2021).
106. V. S. Patel, S. J. Cooper, J. E. Deakin, B. Fulton, T. Graves, W. C. Warren, R. K. Wilson, J. A. Graves, Platypus globin genes and flanking loci suggest a new insertional model for beta-globin evolution in birds and mammals. *BMC Biol.* **6**, 34 (2008).

Acknowledgments: E. Salimova, G. Zheng, and M. de Veer at Monash Biomedical Imaging provided valuable assistance in the collection of MRI data. The support and resources from the Center for High Performance Computing at the University of Utah are gratefully acknowledged. We thank the Melbourne Trace Analysis for Chemical, Earth and Environmental Sciences (TrACEES) Platform for access to the micro-CT scanner. **Funding:** Funding support to A.J.P. and S.R.F. was provided by BioPlatforms Australia and Oz Mammals Genomics, with special thanks to J. Deakin and A. Gilbert. Funding for S.L. and N.C. was provided by the National Eye Institute as part of the NIH (EY030546). The computational resources used were partially funded by the NIH Shared Instrumentation Grant 1S10OD021644-01A1 to N.C. **Author contributions:** S.R.F. and A.J.P. conceived the study and wrote the manuscript with help from all coauthors. S.L. performed analyses of eye genes. C.Y.F., G.H., and P.G. performed genome assembly. B.R.M. performed analysis of testicular descent genes. R.D.P.-I. and S.R.F. performed DNA extractions and Oxford Nanopore sequencing. M.K., A.N., and J.R.B. performed 3D scanning and analyses. L.D., R.S., and J.S. performed retrophylogenomic and MSMC analyses. S.R.F. and S.C. performed analysis of hemoglobin genes. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. *N.B.* sources of images acknowledged above can be checked here: Marsupial mole: https://commons.wikimedia.org/wiki/File:Phylogenetic_tree_of_marsupials_derived_from_retroposon_data_(Notoryctemorphia).png. Lens: https://medialibrary.nei.nih.gov/media/3714. Cone and rod: https://medialibrary.nei.nih.gov/media/3715. Koala: https://www.inaturalist.org/observations/118284563. Kangaroo: https://commons.wikimedia.org/wiki/File:Macropus_giganteus_1_(49535146378).jpg. Wombat: https://www.inaturalist.org/observations/187904740.

Submitted 13 February 2024
Accepted 25 November 2024
Published 1 January 2025
10.1126/sciadv.ado4140