

# Ancestry of the Australian Termitivorous Numbat

Anja Zemann,<sup>1</sup> Gennady Churakov,<sup>1</sup> Stephen Donnellan,<sup>2</sup> Frank Grützner,<sup>3</sup> Fangqing Zhao,<sup>‡,4</sup> Jürgen Brosius,<sup>1</sup> and Jürgen Schmitz<sup>\*,1</sup>

<sup>1</sup>Institute of Experimental Pathology (ZMBE), University of Münster, Münster, Germany

<sup>2</sup>Evolutionary Biology Unit, South Australian Museum, Adelaide, Australia

<sup>3</sup>School of Molecular & Biomedical Science, University of Adelaide, Adelaide, Australia

<sup>4</sup>Center for Comparative Genomics and Bioinformatics, Pennsylvania State University

<sup>‡</sup>Present address: Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing, China

**\*Corresponding author:** E-mail: [jueschm@uni-muenster.de](mailto:jueschm@uni-muenster.de).

**Associate editor:** Emma Teeling

All derived sequences have been deposited in the NCBI database (accession nos. KC619625–KC620125).

## Abstract

The Australian numbat, *Myrmecobius fasciatus*, is the only marsupial that feeds almost exclusively on termites and that has a life following the diurnally restricted and dynamic geographical distribution of termites. The millions of years of this adaptation led to unique morphological and anatomical features, especially basicranial and dental characteristics, that make it difficult to identify a clear phylogenetic affiliation to other marsupials. From DNA sequence analyses, the family Myrmecobiidae is placed within the dasyuromorph marsupials, but the exact position varies from study to study, and support values are mostly rather modest. Here, we report the recovery and analysis of approximately 110,000 quasi-fossilized traces of mobile element insertions into the genome of a dasyurid marsupial (Tasmanian devil), 25 of which are phylogenetically informative for early dasyuromorph evolution. Fourteen of these ancient retroposon insertions are shared by the 16 Dasyuromorphia species analyzed, including the numbat, but are absent in the outgroups. An additional 11 other insertions are present in all Dasyuridae but are absent in the numbat. These findings place numbats as the sister group to all living Dasyuridae and show that the investigated Dasyuromorphia, including the Myrmecobiidae, constitutes a monophyletic group that is separated from Peramelemorphia, Notoryctemorphia, and other marsupials.

**Key words:** retroposon phylogenomics, Dasyuromorphia, Myrmecobiidae, Dasyuridae, numbat.

Approximately 140 extant Australian marsupial species are classified in four ordinal clades. The largest order Diprotodontia comprises kangaroos, wallabies, and possums among others. The order Notoryctemorphia contains just two species of marsupial moles. The order Peramelemorphia includes the omnivorous bandicoots and bilbies. Finally, the order Dasyuromorphia comprises the insectivorous or carnivorous numbat, quolls, dunnarts, Tasmanian devil, and the extinct thylacine or Tasmanian tiger.

The Southern Australian termitivorous numbat (*Myrmecobius fasciatus*) represents the monotypic family Myrmecobiidae (Gill 1872) and is special among marsupials due to its extreme specialization, making it difficult to reconstruct its accurate phylogeny. The first recorded sight of a numbat goes back to 1831 (summarized in Friend 1989). At the beginning of the 20th century, based on somewhat unusual dentition, Bensley (1903) concluded that the numbat was derived possibly from an early dasyurid. Using comparative serology, Kirsch (1968) proposed *Myrmecobius* as a distinct lineage within Dasyuridae, but analyses of mitochondrial *Cytochrome b* and 12S rDNA and protamine nuclear gene sequences did not place *Myrmecobius* separate from the dasyurid lineage but rather indicated some specific affinity to dunnarts (Krajewski et al. 1997). These analyses did,

however, support the monophyly of Dasyuromorphia (numbat, thylacine, and dasyurids). It was concluded that morphological characters are less suitable for supporting Dasyuromorphia monophyly because they are dominated by either ancestral or uniquely derived character states. Moreover, the paucity of fossils further hinders resolution of the controversial evolution of myrmecobiids and dasyurids. With an improved and reanalyzed data set of *Cytochrome b*, 12S rDNA, and protamine gene sequences, Krajewski et al. (2000) did find weak support for a separation of numbat and thylacine from dasyurids. Later studies with greatly improved phylogenetic resolution through the use of multiple nuclear genes unfortunately have not included *Myrmecobius* (Meredith et al. 2008). Even a multiple tree consensus supertree, combining 158 source trees from 107 publications and involving all available character types (morphology, immunology, karyotypes, DNA hybridization, DNA sequences, and behavioral data) did not determine how the thylacines and numbats are related to dasyurids (Cardillo et al. 2004). A study of five nuclear genes from representatives of all extant marsupial families did provide strong support for a basal split between Dasyuridae and Myrmecobiidae (Meredith et al. 2009; see also Meredith et al. 2011). In contrast to mitochondrial single gene reconstructions, the mitochondrial genome sequences of

marsupials, including the ancient mitochondrial DNA of the Tasmanian tiger, revealed additional significant support for Dasyuridae and Dasyuromorphia (Miller et al. 2009).

As retroposon insertion patterns avoid some of the pitfalls inherent in sequence-based data (Schmitz et al. 2001; Churakov et al. 2009), we sought to test these phylogenetic signals with evidence from a totally different, sequence-independent, type of analysis. In this study, we employed molecular “fossils” that are present in the genomes of all mammals and are passed on from lineage to lineage up to the contemporary species. Such ancient molecular relicts arise via transcription and retroposition of genomic sequences and remain recognizable in genomes for hundreds of millions of years. Thus, specific retroposons established in the genome of a common ancestor and subsequently passed on to descendent species leave reliable indicators of their common origin and hence constitute a strong phylogenetic signature. Ancient retroposons can be retrieved in present-day genomes by a combination of bioinformatic analysis, polymerase chain reaction (PCR), and sequencing. We used these unique insertion signatures successfully to reconstruct the complete order-level phylogeny of marsupials (Nilsson et al. 2010), mammals (Kriegs et al. 2006; Warren et al. 2008), and birds (Suh et al. 2011). In our marsupial study, four independent, shared retroposon insertions provided significant evidence confirming the single common origin of the four Australian marsupial orders Dasyuromorphia, Diprotodontia, Notoryctemorphia, and Peramelemorphia.

## Results and Discussion

Here, we searched for potential phylogenetically informative retroposons from early Dasyuromorphia using the genomic sequence information of the Tasmanian devil. To access suitable relevant elements active during early dasyuromorph evolution, we first conducted a Transposition in Transposition (TinT) analysis (Churakov et al. 2010), comparing the Tasmanian devil to wallaby and opossum, based on the fact that young active elements jump into old inactive elements but not vice versa. Such directionally nested retroposons contain complex information about the historical succession of elements, suitable data for finding the ones active at a certain ancient divergence. For these phylogenetic applications, we focused on Short Interspersed Elements (SINEs) because they are frequent and their restricted size facilitates easy amplification and sequencing of the corresponding loci. Analysis of the multidimensional direction of nested insertions showed that WSINE1 (Wallaby Long Interspersed Element [LINE1]-mobilized SINE) and WALLSI (Wallaby retroposon-like transposable element-mobilized SINE) elements were especially active at the branch leading to the Tasmanian devil (fig. 1). From approximately 110,000 analyzed genomic retroposon insertion loci (41,229 WALLSI1/2 loci and 69,008 WSINE1 loci) active during a relatively narrow phylogenetic window early in dasyuromorph evolution, we selected the 161 most conserved loci (13 WALLSI1/2 loci and 148 WSINE1 loci) for comparative PCR amplification and/or in silico sequence analysis in 16 representative Dasyuromorphia species plus eight nondasyuromorph outgroups species (fig. 2). On the

basis of the gel patterns from those that were amplifiable in all species and phylogenetically informative, we experimentally amplified 500 loci and analyzed 573 sequences from 24 loci in the species in figure 2 to verify the orthologous presence or absence of 25 diagnostic phylogenetic markers of the early dasyuromorph diversification. Fourteen of these are present in all 16 Dasyuromorphia species including the numbat and are absent in other marsupials; a further 11 markers are present in all Dasyuridae but are absent in the numbat (fig. 2). This significant number of unique rare genomic changes enables us to assign the numbats as the sister group to all living Dasyuridae. Together, Dasyuridae and numbats build a monophyletic group that is clearly separated from Peramelemorphia and Notoryctemorphia, their next natural sister groups.

These data resolve previously controversial or weak reconstructions (Krajewski et al. 1997, 2000) and confirm the findings of the nuclear sequence study by Meredith et al. (2009) and the position of the numbat in the study of mitochondrial genomes (Miller et al. 2009) with sequence-independent evidence. The present data set will be especially suitable for testing the phylogenetic position of the extinct marsupial wolf; whether the thylacine split from the dasyurids before or after *Myrmecobius* diverged from the dasyuromorph lineage can be determined by examining the presence/absence patterns of these diagnostic retroposon loci in the thylacine. However, current biological material and available genomic sequence information of the extinct thylacine are not suitable for reconstructing the large diagnostic retroposon loci (>800 nt), and ancient DNA analyses are restricted to small nuclear amplicates (usually much smaller than 200 nucleotides, Austin J, personal communication).

The uniqueness of *Myrmecobius*, manifested not only in their exclusive eating habits but also in their isolated phylogenetic position distinct from dasyurids, is now substantiated with a significant number of phylogenetically informative markers supporting its inclusion in Dasyuromorphia (14 SINE markers) and its exclusion from Dasyuridae (11 SINE markers) (fig. 2). Intensive effort to protect these animals is now the next and important step to retain a unique species along with the broad diversity of marsupials.

## Materials and Methods

### Marsupial Material

All necessary samples for experimental PCR amplification were provided by the South Australian Museum in Adelaide (ABTC voucher number), Australia, or the DNA-Zoo collection of Professor Harald Jockusch (HJ).

Diprotodontia: *Trichosurus vulpecula* (Tvu-HJ).

Notoryctemorphia: *Notoryctes typhlops* (Nty-ABTC27520).

Peramelemorphia: *Perameles bougainville* (Pbo-ABTC10622),

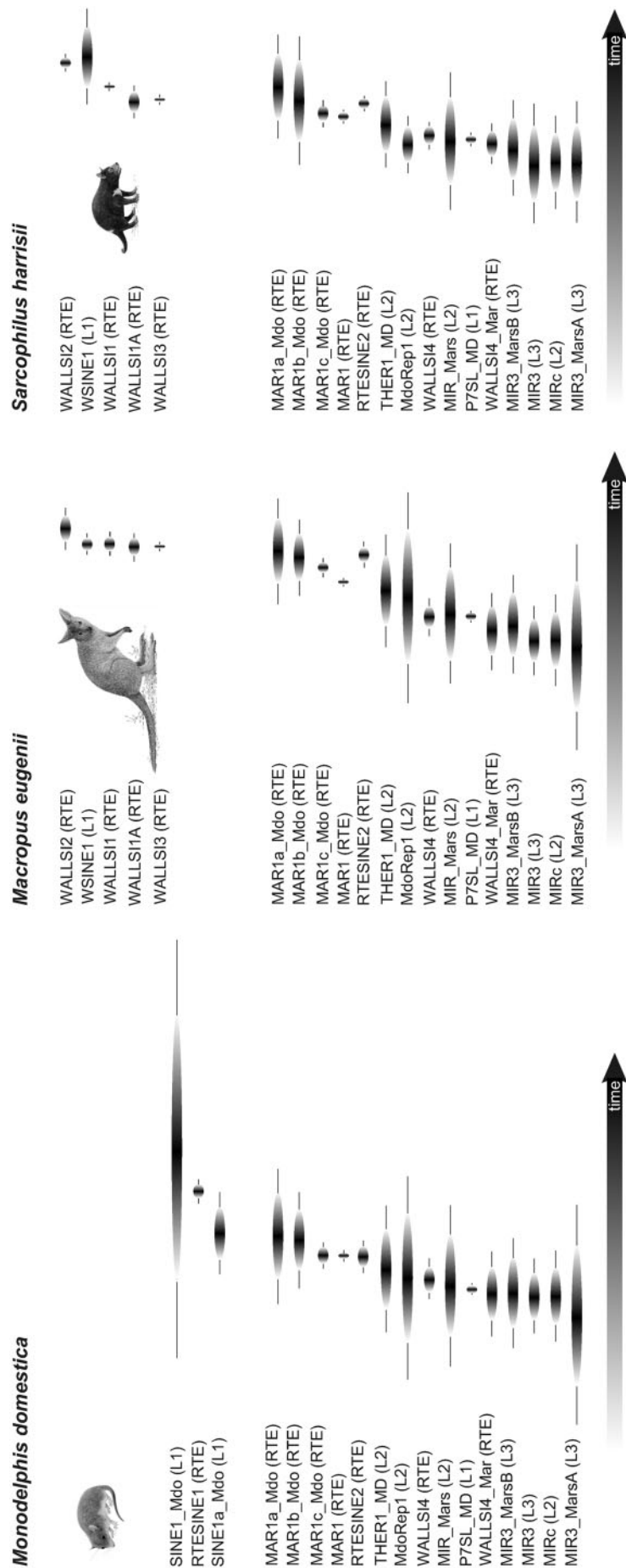
*Echymipera rufescens* (Eru-ABTC48922), *Isoodon obesulus* (Iob-ABTC83719), and *Macrotis lagotis* (Mla-ABTC27581).

Dasyuromorphia: *Myrmecobius fasciatus* (Mfa-ABTC104467),











*Sminthopsis macroura* (Sma-ABTC00539), *Planigale gelesi*

(Pge-ABTC00545), *Antechinus stuartii* (Ast-ABTC01044),

*Murexia longicauda* (Mlo-ABTC07563), *Phascogale*



**Fig. 1.** TinT analysis of opossum (*Monodelphis domestica*), wallaby (*Macropus eugenii*), and Tasmanian devil (*Sarcophilus harrisi*) SINEs. Ovals represent SINE TinTs retroposed by LINE3 (L3), LINE2 (L2), LINE1 (L1), and retroposon-like transposable elements (RTE). The 50% probability of the activity distribution is shown by ovals and the 90% probability of activity distribution by horizontal lines. The relative time axes are shown at the bottom. Shared marsupial retroposon activities are shown at the lower part of the figure. All abbreviations of SINEs are taken from the giji database (<http://www.gijinst.org/> version from 20 September 2011).

		Marker:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Didelphimorphia																											
	<i>Monodelphis domestica</i>		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Diprotodontia																											
	<i>Macropus eugenii</i>		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	<i>Trichosurus vulpecula</i>		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Notoryctemorphia																											
	<i>Notoryctes typhlops</i>		---	---	---	---	---	d	---	---	---	---	---	d	?	---	---	---	---	---	---	---	---	---	---	?	---
Peramelemorphia																											
	<i>Perameles bougainville</i>		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	<i>Echymipera rufescens</i>		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	<i>Isodon obesulus</i>		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	<i>Macrotis lagotis</i>		---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Dasyuromorphia																											
	<i>Myrmecobius fasciatus</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Dasyuridae																											
	<i>Sminthopsis macroura</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Planigale gilesi</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	?	+	+
	<i>Antechinus stuartii</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Murexia longicaudata</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Phascogale tapoatafa</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Myoictis leucura</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Dasyuroides byrnei</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Dasyercus cristicauda</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Pseudantechinus macdonnellensis</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Dasykaluta rosamondae</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Parantechinus apicalis</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Parantechinus bilarni</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Phascosorex dorsalis</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Dasyurus geoffroii</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	<i>Sarcophilus harrisii</i>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

**FIG. 2.** Presence/absence patterns of retroposon insertions. In addition to the information from the three available genomes of *Monodelphis*, *Macropus*, and *Sarcophilus*, a total of 24 loci were amplified by PCR from the remaining 21 species and examined for the presence (+) or absence (–) of the 25 diagnostic retroposon insertions (markers 7 and 20 are from the same locus). A (?) indicates the failure of amplification and (d) a deletion of the insertion locus. The marker numbers at the top correspond to those of the primers in [supplementary table S1, Supplementary Material](#) online. Fourteen markers support the monophyly of Dasyuromorphia and an additional 11 markers the monophyly of the investigated Dasyuridae.



*tapoatafa* (Pta-ABTC27600), *Myoictis leucura* (Mle-ABTC45347), *Dasyuroides byrnei* (Dby-ABTC07555), *Dasyercus cristicauda* (Dcr-ABTC37971), *Pseudantechinus macdonnellensis* (Pma-ABTC91679), *Dasykaluta rosamondae* (Dro-ABTC07549), *Parantechinus apicalis* (Pap-ABTC07518), *P. bilarni* (Pbi-ABTC29664), *Phascosorex dorsalis* (Pdo-ABTC42508), and *D. geoffroi* (Dge-ABTC07519).

Additional genome sequences: *Monodelphis domestica* (<http://hgdownload.cse.ucsc.edu/downloads.html#opossum>), *Macropus eugenii* (<http://hgdownload.cse.ucsc.edu/downloads.html#wallaby>), and *Sarcophilus harrisii* ([http://hgdownload.cse.ucsc.edu/downloads.html#tasmanian\\_devil](http://hgdownload.cse.ucsc.edu/downloads.html#tasmanian_devil)).

### TinT Analysis

To derive the relative periods of activity for the different marsupial SINE elements, we performed complete genome repeat masking using a local version of the RepeatMasker (<http://www.repeatmasker.org/RMDownload.html>). We used the RepeatMasker library from the Genetic Information Research Institute (giri; <http://www.girinst.org/> version from 20 September 2011). The corresponding output files were the source for the TinT application (<http://www.compugen.uni-muenster.de/tools>; Churakov et al. 2010) with default parameters for SINEs. Marsupial-specific SINEs were selected to calculate the TinT graphs shown in figure 1.

### Screening for Potential Phylogenetically Informative WSINE1 and WALLS1/2

We extracted more than 69,000 WSINE1 and 41,000 WALLS1/2 loci from the Tasmanian devil plus 700 nt of flanking sequences for each locus. These sequences were projected (via Basic Local Alignment Search Tool search) against a compilation of exon–intron–exon units from opossum and wallaby. Only those loci exhibiting the presence of intronic WSINE1 or WALLS1/2 insertions in the Tasmanian devil and their absence in opossum and wallaby were selected for further investigations. A total of 161 such loci (148 WSINE1, 6 WALLS1, and 7 WALLS2) with intron-embedded SINEs were selected to generate highly conserved exonic PCR primers for PCR amplification and experimental detection of the presence/absence boundary in the listed marsupial species (see earlier). This search strategy corresponds to strategy II in Suh et al. (2011).

### PCR Amplification and Sequencing of Diagnostic Loci

We followed standard protocols for PCR amplification, cloning, and sequencing of the selected genomic loci (Kriegs et al. 2006) using the PCR primers presented in [supplementary table S1, Supplementary Material](#) online. From approximately 110,000 computationally selected genomic loci, we selected 161 for PCR amplification and gel electrophoresis. After inspecting the gel images for diagnostic insertions exhibiting size changes of approximately 120–300 nt for Dasyuridae or Dasyuridae plus *Myrmecobius*, we sequenced 24 of these loci (locus 7 contains markers 7 and 20).

## Supplementary Material

Supplementary table S1 is available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

## Acknowledgments

The authors thank Jón Baldur Hlíðberg for providing the marsupial paintings; Marsha Bundman for editorial help; Leanne Wheaton for organizing the sample transfer from Australia to Germany; Webb Miller and Stephan Schuster from Pennsylvania State University for an early access to the Tasmanian devil genome; and Harald Jockusch for providing additional samples. This work was supported by the Deutsche Forschungsgemeinschaft (SCHM1469/3-1, SCHM1469/3-2).

## References

- Bensley BA. 1903. On the evolution of the Australian Marsupialia; with remarks on the relationships of the marsupials in general. *Trans Linn Soc Lond.* 9:83–217.
- Cardillo M, Bininda-Emonds ORP, Boakes E, Purvis A. 2004. A species-level phylogenetic supertree of marsupials. *J Zool Lond.* 264:11–31.
- Churakov G, Grundmann N, Kuritzin A, Brosius J, Makalowski W, Schmitz J. 2010. A novel web-based TinT application and the chronology of the primate *Alu* retroposon activity. *BMC Evol Biol.* 10:376.
- Churakov G, Kriegs JO, Baertsch R, Zemmann A, Brosius J, Schmitz J. 2009. Mosaic retroposon insertion patterns in placental mammals. *Genome Res.* 19:868–875.
- Friend JA. 1989. *Myrmecobiidae*. In: Walton DW, Richardson BJ, editors. *Fauna of Australia, Mammalia*. Canberra (Australia): Australian Government Publishing Service. p. 583–590.
- Gill T. 1872. Arrangement of the families of mammals with analytical tables. *Smithson Misc Collect.* 11:1–98.
- Kirsch JAW. 1968. Prodrum of the comparative serology of Marsupialia. *Nature* 217:418–420.
- Krajewski C, Buckley L, Westerman M. 1997. DNA phylogeny of the marsupial wolf resolved. *Proc Biol Sci.* 264:911–917.
- Krajewski C, Wroe S, Westerman M. 2000. Molecular evidence for the pattern and timing of cladogenesis in dasyurid marsupials. *Zool J Linn Soc.* 130:375–404.
- Kriegs JO, Churakov G, Kiefmann M, Jordan U, Brosius J, Schmitz J. 2006. Retroposed elements as archives for the evolutionary history of placental mammals. *PLoS Biol.* 4:e91.
- Meredith RW, Janecka JE, Gatesy J, et al. (22 co-authors). 2011. Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science* 334:521–524.
- Meredith RW, Krajewski C, Westerman M, Springer MS. 2009. Relationships and divergence times among the orders and families of Marsupialia. *Bull Museum Northern Arizona* 65:383–406.
- Meredith RW, Westerman M, Springer MS. 2008. A timescale and phylogeny for “bandicoots” (Peramelemorphia: Marsupialia) based on sequences for five nuclear genes. *Mol Phy Evol.* 47:1–20.
- Miller W, Drautz DI, Janecka JE, et al. (20 co-authors). 2009. The mitochondrial genome sequence of the Tasmanian tiger (*Thylacinus cynocephalus*). *Genome Res.* 19:213–220.
- Nilsson MA, Churakov G, Sommer M, Tran NV, Zemmann A, Brosius J, Schmitz J. 2010. Tracking marsupial evolution using archaic genomic retroposon insertions. *PLoS Biol.* 8:e1000436.
- Schmitz J, Ohme M, Zischler H. 2001. SINE insertions in cladistic analyses and the phylogenetic affiliations of *Tarsius bancanus* to other primates. *Genetics* 157:777–784.
- Suh A, Paus M, Kiefmann M, Churakov G, Franke FA, Brosius J, Kriegs JO, Schmitz J. 2011. Mesozoic retroposons reveal parrots as the closest living relatives of passerine birds. *Nat Commun.* 2:443.
- Warren WC, Hillier LW, Graves JAM, et al. (104 co-authors). 2008. Genome analysis of the platypus reveals unique signatures of evolution. *Nature* 453:175–183.