Disentangling the relationship of the Australian marsupial orders using retrotransposon and evolutionary network analyses

Susanne Gallus[1], Axel Janke[1,2], Vikas Kumar[1], Maria A. Nilsson[1]*

[1]Senckenberg Biodiversity and Climate Research Centre, Senckenberg Gesellschaft für Naturforschung, Senckenberganlage 25, D-60325 Frankfurt am Main, Germany.

[2]Goethe University Frankfurt Institute for Ecology, Evolution & Diversity Biologicum Max-von-Laue-Str.13, D-60439 Frankfurt am Main, Germany.

*Author for correspondance: Dr. Maria Nilsson

Senckenberg - Biodiversity and Climate Research Centre, BiK-F

Senckenberganlage 25 Frankfurt am Main, Germany Tel: +49 69-75421829

email: maria.nilsson-janke@senckenberg.de

Abstract

The ancestors to the Australian marsupials entered Australia around 60 (54-72) million years ago from Antarctica, and radiated into the four living orders Peramelemorphia, Dasyuromorphia, Diprotodontia and Notoryctemorphia. The relationship between the four Australian marsupial orders has been a long-standing question, because different phylogenetic studies were not able to consistently reconstruct the same topology. Initial *in silico* analysis of the Tasmanian devil genome and experimental screening in the seven marsupial orders revealed 20 informative transposable element insertions for resolving the inter- and intraordinal relationships of Australian and South American orders. However, the retrotransposon insertions support three conflicting topologies regarding Peramelemorphia, Dasyuromorphia and Notoryctemorphia, indicating that the split between the three orders may be best understood as a network. This finding is supported by a phylogenetic re-analysis of nuclear gene sequences, using a consensus network approach that allows depicting hidden phylogenetic conflict, otherwise lost when forcing the data into a bifurcating tree. The consensus network analysis agrees with the transposable element analysis in that all possible topologies regarding Peramelemorphia, Dasyuromorphia, and Notoryctemorphia in a rooted four-taxon topology are equally well supported. In addition, retrotransposon insertion data supports the South American order Didelphimorphia being the sistergroup to all other living marsupial orders. The four Australian orders originated within three million years at the Cretaceous-Paleogene boundary. The rapid divergences left conflicting phylogenetic information in the genome possibly generated by incomplete lineage sorting or introgressive hybridisation, leaving the relationship among Australian marsupial orders unresolvable as a bifurcating process million years later.

Introduction

Much effort has been placed into resolving the relationship among the more than 300 living marsupial species (Szalay 1982; Kirsch et al. 1997; Nowak 2005; Nilsson et al. 2010; Meredith et al. 2008, 2011; Mitchell et al. 2014). The most dramatic revelation was when a small South American marsupial was taxonomically moved from within the order Didelphimorphia to its own monotypic order, Dromiciopsa (later: Microbiotheria) (Szalay 1982). A single morphological character suggested that Microbiotheria might be closely related to Australian marsupials (Szalay 1982). Molecular analyses have since identified its position as sistergroup to all Australian marsupials (Nilsson et al. 2010; Meredith et al. 2011). However, the relationship among the four Australian marsupial orders (Peramelemorphia, Dasyuromorphia, Diprotodontia and Notoryctemorphia) is still highly controversial and has yet to be settled (Nilsson et al. 2010; Meredith et al. 2011; Mitchell et al. 2014). In particular the phylogenetic position of Notoryctemorphia (marsupial moles) is debated because morphological and molecular approaches result in conflicting topologies or non-significant branch support (Supplementary table 1). Most molecular studies favor placing Peramelemorphia, Dasyuromorphia and Notoryctemorphia in conflicting constellations to the exclusion of Diprotodontia. Notoryctemorphia consist of one living genus, *Notoryctes*, which contains two species, the northern (*N. caurinus*) and the southern marsupial mole (*N. typhlops*). Marsupial moles resemble the placental moles (e.g. golden moles, Chrysochloridae) regarding their convergent adaptations to the sub-terrestrial, fossorial lifestyle (Kirsch et al. 1997; Nowak 2005; Archer et al. 2011). The phenotypic adaptation has led to highly derived characters that complicate morphological phylogenetic analyses of the position of Notoryctemorphia (Horovitz and Sánchez-Villagra 2003; Asher et al. 2004).

The analysis of presence and absence of retrotransposon insertions in mammalian genomes is a powerful tool to resolve conflicting phylogenetic hypotheses (Shedlock and Okada 2000;

Boore 2006) and to recover evolutionary processes that would remain hidden by nucleotide sequence analyses (Terai et al. 2003). The availability of whole genome sequence data has revolutionized the use of retrotransposon insertion approaches for phylogenetic questions by increasing the amount of investigable loci (Warren et al. 2008; Nishihara et al. 2009). Retrotransposons propagate from a few active 'source elements' in the genome following the 'copy-and-paste' principle using RNA-intermediates (Kramerov and Vassetzky 2011). Once new retrotransposon copies are inserted in the genome they are usually unable to propagate further (Shedlock and Okada 2000). With the knowledge that insertions that occurred in the common ancestor will be transferred to the offspring in the same orthologous position, the fossilized retrotransposon insertions in the mammalian genome are ideal to entangle the relationship among species (Shedlock and Okada 2000).

Retrotransposon insertion analyses are less affected by the stochastic processes that can confound sequence-based studies (Shedlock and Okada 2000; Ray et al. 2006). Nucleotide composition biases, varying evolutionary rates and parallel/convergent substitutions from limited character states (GATC) cause uncertainties in sequence-based tree reconstruction (Boore 2006). Retrotransposons have the advantage of a clear knowledge of the ancestral state, which is absence (Ray et al. 2006). The probabilities of independent retrotransposon insertions into the same target sequence at the exact same location as well as exact deletions are negligible (van der Lagemaat et al. 2005). Retrotransposon insertion analyses have resolved the relationship among several marsupial and placental mammalian groups (Shimamura et al. 1997; Churakov et al. 2009; Nilsson et al. 2010). The availability of a dasyuromorphian marsupial genome (Tasmanian devil) made it possible to use retrotransposon insertion analysis for investigating the controversial phylogenetic relationships among Australian marsupial orders. An experimental screen of 158 *in silico* identified introns with suitable retrotransposon insertions revealed 23 phylogenetically

informative insertions for resolving intra- and interordinal relationships in both South American and Australian marsupial orders (Supplementary table 2).

The South American Didelphimorphia is the sistergroup to all marsupials

The earliest fossil crown group marsupials bear morphological similarities to the South American order Didelphimorphia (opossums) that has been regarded as sistergroup to all other living marsupial orders (Horovitz and Sánchez-Villagra 2003; Horovitz et al. 2009). However, some sequence-based phylogenetic studies place the South American order Paucituberculata (shrew opossums) as the sister group to living marsupials (Meredith et al. 2011; Mitchell et al. 2014). The apparent conflict between sequence data and the fossil record, that favor two different marsupial orders at the basal position of all living marsupials, has so far not been further investigated (Meredith et al. 2011; Mitchell et al. 2014).

To obtain statistically significant support for a branch, three retrotransposon insertions are required (Waddell et al. 2001). We identified three retrotransposon insertions (MIR3, Mar1_MD, MIRc) present in all living marsupial orders supporting the monophyly of marsupials (Figure 1), but none for conflicting topologies, yielding a [3 0 0] pattern of support and a probability for this branch of p= 0.037 (Waddell et al. 2001).

In the course of our screening we found three retrotranposons (WSINE1, WALLSI1, RTESINE2) favoring Didelphimorphia as sister group to all other living marsupials (Figure 1). The two SINE types WALLSI1 and RTESINE2 do not exist in the didelphimorphian genome (Gentles et al. 2007), further strengthening the support for placing Didelphimorphia as the earliest divergence. Previous analysis of retrotransposon insertions identified two SINE insertions ([2 0 0] p= 0.111) supporting Didelphimorphia as sister group to all other living marsupials (Nilsson et al. 2010), resulting in a total of five independent retrotransposon

insertions ([5 0 0] p= 0.001) without any conflicting markers. However, given the discrepancies between the retrotransposon insertions and previous sequence analyses it is still possible that analyses of future genome data, in particular that of Paucituberculata, might result in conflicting data.

Intraordinal relationships among and between Australian marsupials

Within Marsupialia, the order Dasyuromorphia, the carnivorous marsupials, have attracted considerable attention in morphological and molecular phylogenetic studies, because they represent the dominant terrestrial mammalian carnivores of Australia and New Guinea. The order is sub-divided into three families (Archer 1984). They include the extinct Tasmanian wolf (*Thylacinus*), the endangered numbat and the species-rich family Dasyuridae. The latter includes 69 species such as quolls, dibblers, dunnarts, and the Tasmanian devil (Nowak 2005). The earliest dasyuromorphian fossils from middle Miocene (23.5-11 million years ago (mya)) represent mainland Australian *Thylacinus* species (Wroe 1999). The ant-eating numbat (Myrmecobiidae) is only known since Holocene (since 10,000 years ago), and only few species belonging to Dasyuridae have been found at Riversleigh, an Oligocene/Miocene fossil site, 25-15 mya (Archer 1984; Long et al 2002). The three Dasyuridae subfamilies are known from Pliocene deposits, 5.3-2.6 mya (Wroe 2003). Given the poor fossil record from Dasyuromorphia, phylogenetic conclusions, divergence times and evolutionary correlations to environmental changes during the Miocene (23.0-5.3 mya) depend largely on molecular data (Krajewski et al. 2000a, b). Three novel retrotransposon insertions further strengthen the divergence between Dasyuridae and Myrmecobiidae ([3 0 0] p= 0.037) and the monophyly of Dasyuromorphia ([3 0 0] p= 0.037) (Figure 1).

The order Peramelemorphia consists of 23 species that are divided into two families, Peramelidae (bandicoots) and Thylacomyidae (bilbies) (Nowak 2005). Thylacomyidae contain just one extant species, the Greater Bilby (*Macrotis lagotis*), as the Lesser Bilby (*Macrotis leucura*) became extinct during the 20th century (Nowak 2005; Gurovich et al. 2014). Peramelemorphian fossils are known from fossil sites of Oligocene/Miocene age (eg. Riversleigh), with the earliest crown-group fossils being Middle Miocene (Gurovich et al. 2014, Travouillon et al. 2014). Previous molecular phylogenetic analysis and divergence time estimation placed the age for the basal split within Peramelemorphia at 27 mya (Westerman et al. 2012). We identified four retrotransposon insertions ([4 0 0] p= 0.012) that support the monophyly of Peramelinae.

The relationships inside Peramelemorphia and Dasyuromorphia have been extensively investigated using nuclear genes, complete mitochondrial genomes as well as retrotransposons (Krajewski et al. 2000a, b; Westerman et al. 2012; Meredith et al. 2011; Zeman et al. 2013; Gallus et al. 2014; Mitchell et al. 2014). The results of our retrotransposon insertion analysis are statistically significant and in agreement with previous studies.

Conflicting signals for the position of Notoryctemorphia among Australian marsupial orders

We identified four phylogenetically informative, but conflicting, retrotransposon insertions for the position of Notoryctemorphia, supporting three different topologies ([2 1 1] p= 0.284) (Figure 2). Two WSINE1 insertions support a sistergroup relationship between Dasyuromorphia and Notoryctemorphia that is most often recovered in sequence-based studies (Supplementary table 1). However, one independent insertion supports a relationship between Notoryctemorphia and Peramelemorphia, previously suggested based on syndactyly, a morphological character possibly in common to both orders (Szalay 1982). Syndactyly, the

joining of the third and fourth digit on the hind foot, occurs in two unrelated Australian orders, Diprotodontia and Peramelemorphia, but its presence in Notoryctemorphia is debated (Weisbecker and Nilsson 2008). Finally one retrotransposon marker was identified for the third possible topology in a rooted four-taxon tree. This marker supports a sistergroup relationship between Peramelemorphia and Dasyuromorphia (Figure 2), a relationship in accordance with nuclear sequence data (Meredith et al. 2011). A recent study combining a novel mitogenomic and a nuclear gene dataset (Meredith et al. 2011) recovered the same topology, but with less support (Mitchell et al. 2014).

The strongest support in our analysis came from two retrotransposon insertions that agree with the topology from most sequence-based studies, showing Dasyuromorphia as sistergroup to Notoryctemorphia (Supplementary table 1). The pattern of retrotransposon insertions, however, indicates that there is conflict in the genome and mirrors the inconsistent results based on previous sequence data analyses (Supplementary table 1). The four Australian marsupial orders diverged in a rapid succession close to the Cretaceous-Paleogene (KPg) boundary around 67-64 mya (Mitchell et al. 2014). During the relatively short time spanning the divergences, only a limited amount of mutations could accumulate. In addition, new mutations occur continuously in the species genome, creating phylogenetic noise (Leaché et al. 2014). Incomplete lineage sorting (ILS), a phenomenon where allelic variation does not have time to sort during speciation can lead to problems during phylogenetic reconstruction (Hudson 1983; Tajima 1983; Pinho and Hey 2010,). Another issue is that introgressive hybridization or gene flow might have taken place between the diverging marsupial populations (Hudson 1983; Tajima 1983; Pinho and Hey 2010; Phillips et al. 2013). It is difficult to assess if, and how much ILS and gene flow have contributed to the observed phylogenetic conflict (Pinho and Hey 2010). As neither genome data for Peramelemorphia nor Notoryctemorphia is currently available it is not possible to do an *in silico* whole genome

screen to find phylogenetically informative retrotransposon insertions. Thus, we caution that when complete genome data is available for Peramelemorphia or Notoryctemorphia allowing for a more exhaustive marker screen, it could reach other conclusions, as we can not rule out the possibility of exact deletion or precise insertions of elements in our markers (van de Lagemaat et al. 2005).

Previous studies showed that some nodes in the placental mammalian tree have experienced ILS and possible ancestral gene flow and are impossible to resolve using retrotransposons or sequence-based phylogenetic analyses. Such patterns are observed when the radiation between species is within 2-4 million years, a time frame in which lineage sorting of alleles may not be completed or hybridization between species is still possible (Hallström and Janke 2010). The deepest splits in the placental mammalian tree have been particularly problematic to resolve (Churakov et al. 2009; Nishihara et al. 2009; Hallström and Janke 2010). Retrotransposon insertion analyses have revealed conflicting signals for relationships among Cetartiodactyla, Carnivora, Perissodactyla, and Chiroptera (Nishihara et al. 2006; Hallström et al. 2011), which were not resolvable as a bifurcating tree by phylogenomic analysis (Hallström et al. 2011). It is highly unlikely that the observed signals of ILS in placental mammals are caused by erosion of the retrotransposons, as this would require erosion in all species and orders at the same rate. Or vice versa, meaning that all species within the placental mammalian orders acquired elements independently at the exact same position.

The results from our retrotransposon screen suggest that aside from a possibility of lack of data, it might be not possible to find a single sistergroup to Notoryctemorphia due to the short divergence times between the Australian marsupial orders. Thus we conclude, that some ordinal relationships might better be depicted as an evolutionary network rather than a bifurcating tree. Networks allow the depiction of conflicting signals from phylogenetic trees that are otherwise ignored when striving for a strict bifurcating species tree (Bapteste et al.

2013). Therefore, we re-investigated a published dataset of marsupial nuclear genes (Meredith et al. 2011) for conflicting signals (Figure 3). The re-analysis shows that there is a strong signal for grouping Dasyuromorphia, Peramelemorphia and Notoryctemorphia to the exclusion of Diprotodontia, which has been suggested by sequence-based analyses as well as one retrotransposon marker (Nilsson et al. 2010). However, the consensus network analysis suggests that there is equally strong phylogenetic signal for each of the three possible topologies among Dasyuromorphia, Peramelemorphia and Notoryctemorphia (Figure 3). High levels of conflicting signals in the sequence data are in addition found inside two orders, the Australian Diprotodontia and the South American Didelphimorphia.

The consensus-network analysis coupled with the findings from the retrotransposon insertions suggest that the evolution of the Australian orders was complex and cannot be depicted as a bifurcating tree. The root of the marsupial tree might resemble the basal splits in the placental mammalian phylogenetic tree, which seem to have undergone extensive incomplete lineage sorting as suggested by the analysis of several complete genomes (Churakov et al. 2009; Nishihara et al. 2009; Hallström and Janke 2008; 2010). It is tempting to interpret the occurrence of syndactyly in different orders, as a consequence of the very rapid divergence. Syndactyly could possibly show that there was gene flow between the diverging populations as it is only found in two unrelated orders. In the absence of genome data we identified retrotransposon markers that weakly support three different topologies for the relationship between Dasyuromorphia, Peramelemorphia and Notoryctemorphia. The phylogenetic finding neatly summarizes the efforts of previous sequence based studies that could not agree on the relationship of the Australian orders. The divergence between the four Australian orders took place in a window of a few million years at the Cretaceous-Paleogene boundary (Mitchell et al. 2014) hindering the accumulation of sufficient phylogenetic information to yield a clear species tree around 60 million years later.

Material and Methods

Phylogenetically informative retrotransposon insertion analysis

We continued the work on an existing dataset comprising 158 marsupial intron markers (Gallus et al. 2015) that were identified by extracting single-copy introns from the Tasmanian devil genome (WTSI_Devil_ref v7.0) (Murchison et al. 2012) and aligning them with homologous loci in the South Amerian opossum (Didelphimorphia) (Mikkelsen et al. 2007) and Tammar wallaby (Diprotodontia) (Renfree et al. 2011). We selected SINEs that were, based on previous studies, expected to be active at the time when the four Australian orders (Diprotodontia, Notoryctemorphia, Peramelemorphia, Dasyuromorphia) diverged (Nilsson et al. 2010; Gallus et al. 2014). In order to avoid a bias in the transposable element selection process, 38 introns without retrotransposable elements were selected by random, to increase the chance of finding independent insertions in species not covered by genome data.

Primers were located in conserved exonic regions flanking the introns with the selected SINE insertions (Supplementary table 3). All introns were amplified and sequenced in representatives from the respective marsupial orders (Supplementary table 4). The Amplicon Taq containing VWR Master Mix (VWR International GmbH, Darmstadt, Germany) was used in a touch-down PCR approach to amplify selected loci. The resulting PCR fragment sizes were inspected on agarose gels for size differences. Where size-shifts between species were observed the locus was Sanger-sequenced in the relevant species (ABI 3730 DNA Analyzer with BigDye terminator sequencing kit 3.1, Applied Biosystems). If direct sequencing of the PCR product was not possible, the PCR-products were cloned into pDrive (Qiagen) or TOPO-TA (Life Technologies) vectors according to the manufacturers recommendations before Sanger sequencing. Sequences were masked with the online tools CENSOR (Kohany et al. 2006) or Repeatmasker (Smit et al. 1996). Sequences for each locus were aligned in Geneious V6 (BioMatters) using the embedded alignment tool, and manually

inspected, and corrected. All alignments were visually screened for retrotransposon type, orientation of retrotransposon, target site duplications and exon/intron homology to ensure the homology of each SINE insertion. The statistical evaluation of SINE insertions was performed according to Waddell et al. (2001) where three insertions per branch are required to obtain significant support [p= 0.037]. The approach calculates branch support under a standard Wright-Fisher coalescent model with panmictic mating, non-overlapping generations and a constant population size (Waddell et al. 2001).

Consensus Network Analyses

Marsupial nuclear gene sequences were downloaded from GenBank (NCBI), the ensembl browser (EMBL-EBI and Wellcome-Trust Sanger institute) and extracted from the Tasmanian devil, opossum and wallaby genome using BLAT (Kent 2002) (Supplementary table 5). In total sequences of 25 marsupial species (Supplementary table 4) were aligned for 26 genes (28 gene fragments, Supplementary table 6), using the MUSCLE (Edgar 2004) algorithm in Geneious (v6) (BioMatters). The Recombination Activation Gene 1 (RAG-1) was split in two smaller fragments, as was Apolipoprotein B (ApoB) due to large amounts of missing sequence information in some species.

All sequence alignments were trimmed to minimize the total amount of gaps and ambiguous nucleotides (nt) in the beginning and end of each alignment. Alignments were visually inspected and corrected if necessary. Best fitting DNA substitution models were calculated for each alignment using the jModelTest (Darriba et al. 2012) with Γ models with four rate categories and a proportion of invariable sites (Supplementary table 7). We used the next complicated substitution model (in most cases GTR), if the proposed model from jModelTest was not implemented in the MrBayes.

Bayesian gene trees were created using MrBayes V3.2.2 for each of the 28 gene fragments (Ronquist and Huelsenbeck 2003). MrBayes was run for 4,000,000 generations using random starting trees with default priors. The burn in was set at 250 (25 % of 1,000 generations). Consensus networks were constructed in SplitsTree4 (version 4.13.1) (Huson and Bryant 2006), using a threshold value of 7% for two different data sets, one consisting of all 28 gene trees, and one data set excluding all trees (n=8) based on alignments shorter than 600 nt (Figure 3, Supplementary figure 1). At this threshold level, multifurcations due to unresolved branches by single genes are not shown, making the network easier to interpret. Additionally, the single genes were concatenated to a dataset consisting of 32,253 nt. The dataset was partitioned to allow different substitution models for each fragment. MrBayes was run using the same parameters on the concatenated dataset as for the single gene dataset (Supplementary figure 2), showing no differences to the consesnsus networks for the main marsupial groups.

Acknowledgements

Figure legends

Figure 1: Phylogenetic tree of the seven marsupial orders based on 16 of the 20 phylogenetically informative retrotransposon insertions and sequence data. Markers for different nodes in the marsupial tree are shown as orange circles. Branches showing retrotransposon support based on Nilsson et al. (2010) are indicated as grey circles. The tree includes representatives of all seven living marsupial orders, and has been scaled to divergence time (Mitchell et al. 2014). The phylogenetic tree is based on a Bayesian analysis of 32,253 nt from 28 nuclear gene fragments (Supplementary figure 1). The branches leading to Peramelemorphia, Dasyuromorphia and Notoryctemorphia have been collapsed. Additional markers supporting the relationship between Peramelemorphia (red shaded area), Dasyuromorphia (green shaded area) and Notoryctemorphia (brown shaded area) are shown separately in Figure 2. Australidelphia refers to the grouping of the South American Microbiotheria and the four Australian orders (Szalay 1982). Q*: Quaternary period; mya: million years ago.

Figure 2: The relationship between Dasyuromorphia, Peramelemorphia and Notoryctemorphia. The retrotransposon screen revealed four informative retrotransposon markers that weakly support three different topologies of the three orders. Retrotransposon support for each topology is indicated by orange colored circles representing the identified number of retrotransposon markers. The small black circles represent markers for intra-ordinal relationships inside the respective orders (Figure 1). Das: Dasyuromorphia (Tasmanian devil); Not: Notoryctemorphia (marsupial mole); Per: Peramelemorphia (Eastern barred bandicoot/ Southern brown bandicoot). Animal figures were provided by Jon Baldur Hlíðberg (www.fauna.is).

Figure 3: Consensus network based on 20 individual Bayesian gene trees (> 600 nt) from 25 marsupial species representing all seven orders. The consensus network (threshold 0.07) shows the possible phylogenetic splits. Red lines and the grey shaded circle highlight the trifurcation between Dasyuromorphia, Notoryctemorphia and Peramelemorphia.

References

Asher RJ, Horovitz I, Sánchez-Villagra MR. 2004. First combined cladistics analysis of marsupial mammal interrelationships. Mol Phylogenet Evol. 33: 240-250.

Archer M 1984. The Australian marsupial radiation. In: Archer M, Clayton G, editors. Vertebrate Zoogeography and Evolution in Australia. Carlisle: Hesperian Press, p. 633–708.

Archer M et al. 2011. Australia's first fossil marsupial mole (Notoryctemorphia) resolves controversies about their evolution and palaeoenvironmental origins. Proc R Soc B. 278: doi:10.1098/rspb.2010.1943.

Bapteste E et al. 2013. Networks: expanding evolutionary thinking. Trends in Genetics. 29:439-441

Boore JL. 2006. The use of genome-level characters for phylogenetic reconstruction. Trends Ecol Evol. 21:439-446.

Churakov G et al. 2009. Mosiac retroposon insertion patterns in placental mammals. Genome Res. 19: 868-875.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 9: 772. doi:10.1038/nmeth.2109

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792-1797.

Gallus S et al. 2015. Evolutionary histories of transposable elements in the genome of the largest living marsupial carnivore, the Tasmanian devil. Mol Biol Evol. In press.

Gentles AJ et al. 2007. Evolutionary dynamics of transposable elements in the short-tailed opossum *Monodelphis domestica*. Genome Res. 17:992-1004.

Gurovich Y, Travouillon KJ, Beck RMD, Muirhead J, Archer M. 2014. Biogeographical implications of a new mouse-sized fossil bandicoot (Marsupialia: Peramelemorphia) occupying a dasyurid-like ecological niche across Australia. J Syst Paleontol. 12:265-290.

Hallström BM, Janke A. 2008. Resolution among major placental mammal interordinal relationships with genome data imply that speciation influenced their earliest radiations. BMC Evol Biol. 8:162. doi:10.1186/1471-2148-8-162.

Hallström BM, Janke A. 2010. Mammalian Evolution May not Be Strictly Bifurcating. Mol Biol Evol. 27:2804-2816.

Hallström BM, Schneider A, Zoller S, Janke A. 2011. A Genomic Approach to Examine the Complex Evolution of Laurasiatherian Mammals. PLoS One. 6:e28199.

Horovitz I, Sánchez-Villagra MR. 2003. A morphological analysis of marsupial mammal higher-level phylogenetic relationships. Cladistics. 19:181-212.

Horovitz I, Martin T, Bloch J, Ladevèze S, Kurz C, Sánchez-Villagra MR. 2009. Cranial anatomy of the earliest marsupials and the origin of opossums. PLoS One. 4:e8278.

Hudson RR. 1983. Testing the constant-rate neutral allele model with protein sequence data. Evolution. 37:203–217.

Huson DH, Bryant D. 2006. Application of Phylogenetic Networks in Evolutionary Studies. Mol Biol Evol. 23:254-267.

Kent WJ. 2002. BLAT – The BLAST-Like Alignment Tool, Genome Res. 12: 656-664

Kirsch JAW, Springer MS, Lapointe F-J. 1997. DNA-hybridization Studies of Marsupials and their Implications for Metatherian Classification. Aust J Zool. 45:211-280.

Kohany O, Gentles AJ, Hankus L, Jurka J. 2006. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinformatics. 25:474. doi:10.1186/1471-2105-7-474.

Krajewski C, Wroe S, Westerman M 2000a. Molecular evidence for the pattern and timing of cladogenesis in dasyurid marsupials. Zoo J of the Linnean Soc. 130:375-404.

Krajewski C, Blacket MJ, Westerman M 2000b. DNA sequence analysis of familial relationships among dasyuromorphian marsupials. J Mammal Evol. 7:95-108.

Kramerov DA, Vassetzky NS. 2011. Origin and evolution of SINEs in eukaryotic genomes. Heredity. 107:487–495. doi:10.1038/hdy.2011.43.

Leaché AD, Harris RB, Rannala B, Yang Z. 2014. The Influence of Gene Flow on Species Tree Estimation: A Simulation Study. Syst Biol. 63:17–30. doi:10.1093/sysbio/syt049.

Long J, Archer M, Flannery T, Hand S. 2002. Prehistoric mammals of Australia and New Guinea: one hundred million years of evolution. Sydney: University of New South Wales Press.

Meredith RW, Westerman M, Case JA, Springer MS. 2008. A Phylogeny and Timescale for Marsupial Evolution Based on Sequences for Five Nuclear Genes. J Mammal Evol. 15:1-36.

Meredith RW et al. 2011. Impacts of the Cretaceous Terrestrial Revolution and KPg Extinction on Mammal Diversification. Science. 334:521-524. doi: 10.1126/science.1211028.

Mikkelsen TS et al. 2007. Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature. 447:167-177. doi:10.1038/nature05805.

Mitchell KJ et al. 2014. Molecular phylogeny, biogeography, and habitat preference evolution of marsupials. Mol Biol Evol. doi:10.1093/molbev/msu176.

Murchison EP et al. 2012. Genome sequencing and analysis of the Tasmanian devil and its transmissible cancer. Cell. 148:780-791.

Nilsson MA et al. 2010. Tracking Marsupial Evolution Using Archaic Genomic Retroposon Insertions. PLoS Biol. 8:e1000436. doi:10.1371/journal.pbio.1000436

Nishihara H, Maruyama S, Okada N. 2009. Retroposon analysis and recent geological data suggest near-simultaneous divergence of the three superorders of mammals. Proc Natl Acad Sci. 106:5235-5240.

Nishihara H, Hasegawa M, Okada N. 2006. Pegasoferae, an unexpected mammalian clade revealed by tracking ancient retroposon insertions. Proc Natl Acad Sci. 103:9929-9934.

Nowak RM. 2005. Walker's marsupials of the world. Baltimore: The Johns Hopkins University Press

Pinho C, Hey J. 2010. Divergence with Gene Flow: Models and Data. Annu Rev Ecol Evol Syst. 41:215-230. doi: 10.1146/annurev-ecolsys-102209-144644.

Phillips MJ, Haouchar D, Pratt RC, Gibb GC, Bunce M. 2013. Inferring Kangaroo Phylogeny from Incongruent Nuclear and Mitochondrial Genes. PloS One. DOI: 10.1371/journal.pone.0057745

Ray DA, Xing J, Salem A-H, Batzer MA. 2006. SINEs of a *Nearly* Perfect Character. Syst Biol. 55:928-935.

Renfree MB et al. 2011. Genome sequence of an Australian kangaroo, *Macropus eugenii*, provides insight into the evolution of mammalian reproduction and development. Genome Biol. 12:R81. doi:10.1186/gb-2011-12-8-r81.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19:1572-1574. doi:10.1093/bioinformatics/btg180.

Shedlock AM, Okada N. 2000. SINE insertions: powerful tools for molecular systematics. Bioessays. 22:148-160.

Smit AFA, Hubley R, Green P. RepeatMasker. http://repeatmasker.org (1996-2014).

Szalay FS. 1982. A new appraisal of marsupial phylogeny and classification. In: Carnivorous Marsupials. Ed: M. Archer. pp. 621-640, Roy Zool Soc New South Wales. Mossman, Australia.

Shimamura M et al. 1997. Molecular evidence from retroposons that whales form a clade within even-toed ungulates. Nature. 388:666-670.

Tajima F. 1983. Evolutionary Relationship of DNA Sequences in Finite Populations. Genetics. 105: 437–460.

Terai Y, Takahashi K, Nishida M, Sato T, Okada N. 2003. Using SINEs to probe ancient explosive speciation: "hidden" radiation of African cichlids? Mol Biol Evol. 20:924-930.

Travouillon KJ, Hand SJ, Archer M, Black KH. 2014. Earliest modern bandicoot and bilby (Marsupialia, Peramelidae and Thylacomyidae) from the Miocene oft he Riversleigh World Heritage Area, northwestern Queensland, Australia. J Vert Paleontol. 34:375-382

Warren WC et al. 2008. Genome analysis of the platypus reveals unique signatures of evolution. Nature. 453:175-183. doi:10.1038/nature06936.

van de Lagemaat LN, Gagnier L, Medstrand P, Mager DL. 2005. Genomic deletions and precise removal of transposable elements mediated by short identical DNA segments in primates. Genome Res. 15:1243-1249. doi: 10.1101/gr.3910705.

Waddell PJ, Kishino H, Ota R. 2001. A phylogenetic foundation for comparative mammalian genomics. Genome Inform Ser Workshop Genome Inform. 12:141-154.

Weisbecker V, Nilsson M. 2008. Integration, heterochrony, and adaptation in pedal digits of syndactylous marsupials. BMC Evol Biol 8:160. doi:10.1186/1471-2148-8-160

Westerman M et al. 2012. Phylogenetic relationships of living and recently extinct bandicoots based on nuclear and mitochondrial DNA sequences. Mol Phylogenet Evol. 62(1):97-108. doi: 10.1016/j.ympev.2011.09.009.

Wroe S, Myers TJ, Wells RT, Gillespie A. 1999. Estimating the weight of the Pleistocene marsupial lion, *Thylacoleo carnifex* (Thylacoleonidae:Marsupialia): implications for the ecomorphology of a marsupial super-predator and hypotheses of impoverishment of Australian marsupial carnivore faunas. Aust J Zool. 47:489-498

Wroe S. 2003. Australian marsupial carnivores: recent advances in palaeontology. In: Jones M, Dickerman C, Archer M, editors. Predators with pouches: the biology of marsupial carnivores. Collingwood: Commonwealth Scientific and Industrial Research Organisation (CSIRO). p. 102–123.

Zemann A et al. 2013. Ancestry of the Australian Termitivorous Numbat. Mol Biol Evol 30:1041-1045. doi: 10.1093/molbev/mst032.

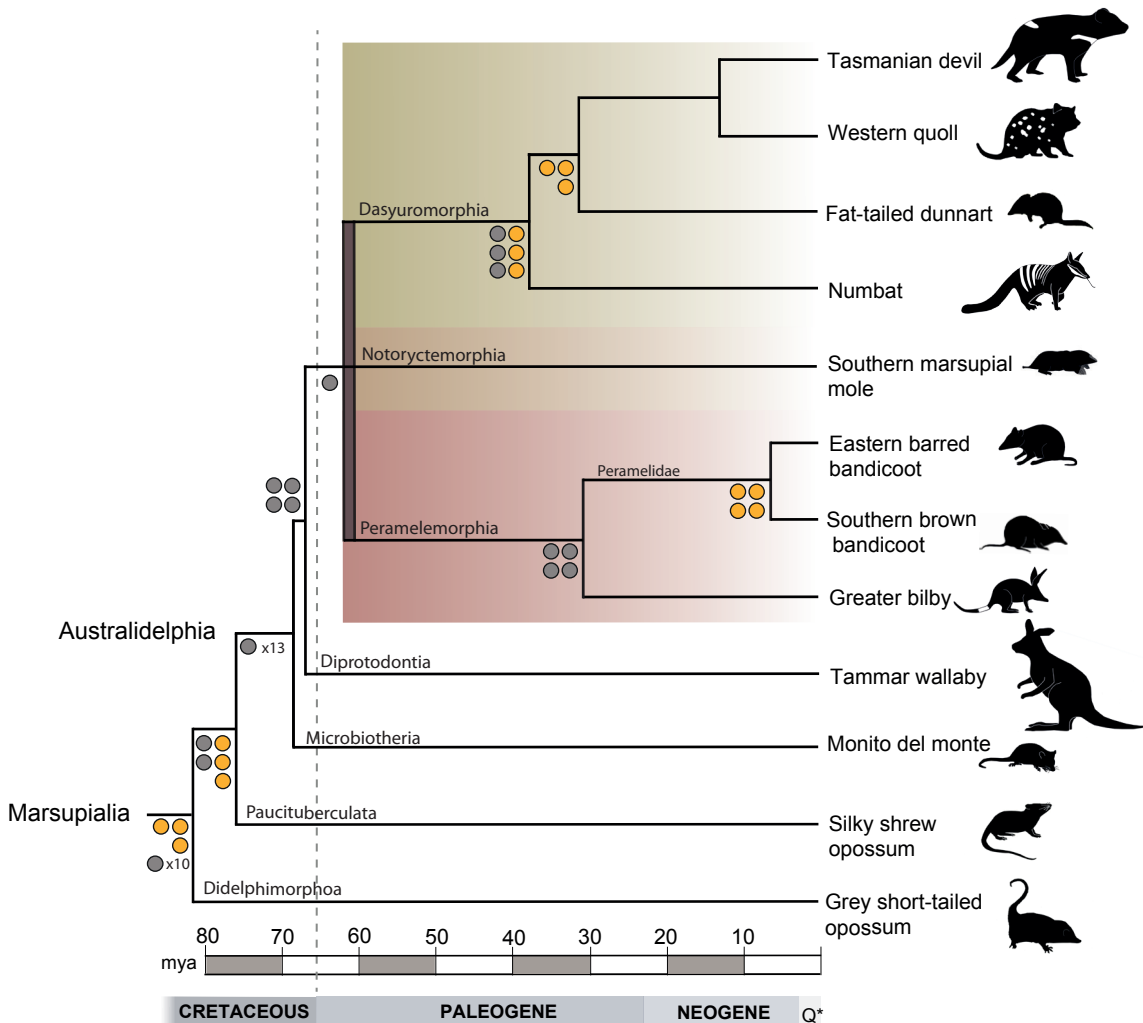Figure 1. Gallus et al.

Figure 2. Gallus et al.

Figure 3. Gallus et al.