Letter

Using genomic data to unravel the root of the placental mammal phylogeny

William J. Murphy,[1,5] Thomas H. Pringle,[2] Tess A. Crider,[1] Mark S. Springer,[3] and Webb Miller[4]

[1]*Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, Texas 77843, USA;* [2]*Sperling Foundation, Eugene, Oregon 97405, USA;* [3]*Department of Biology, University of California, Riverside, California 92521, USA;* [4]*Center for Comparative Genomics and Bioinformatics, Pennsylvania State University, University Park, Pennsylvania 16802, USA*

The phylogeny of placental mammals is a critical framework for choosing future genome sequencing targets and for resolving the ancestral mammalian genome at the nucleotide level. Despite considerable recent progress defining superordinal relationships, several branches remain poorly resolved, including the root of the placental tree. Here we analyzed the genome sequence assemblies of human, armadillo, elephant, and opossum to identify informative coding indels that would serve as rare genomic changes to infer early events in placental mammal phylogeny. We also expanded our species sampling by including sequence data from >30 ongoing genome projects, followed by PCR and sequencing validation of each indel in additional taxa. Our data provide support for a sister-group relationship between Afrotheria and Xenarthra (the Atlantogenata hypothesis), which is in turn the sister-taxon to Boreoeutheria. We failed to recover any indels in support of a basal position for Xenarthra (Epitheria), which is suggested by morphology and a recent retroposon analysis, or a hypothesis with Afrotheria basal (Exafrico-placentalia), which is favored by phylogenetic analysis of large nuclear gene data sets. In addition, we identified two retroposon insertions that also support Atlantogenata and none for the alternative hypotheses. A revised molecular timescale based on these phylogenetic inferences suggests Afrotheria and Xenarthra diverged from other placental mammals ~103 (95–114) million years ago. We discuss the impacts of this topology on earlier phylogenetic reconstructions and repeat-based inferences of phylogeny.

[Supplemental material is available online at www.genome.org. The sequence data from this study have been submitted to GenBank under accession nos. EF122078–EF122139.]

Molecular studies on the interordinal phylogeny of placental mammals have made great progress in the past six years. Large concatenated data sets from both nuclear and mitochondrial genes generally agree on the division of placental mammals into four principal clades, Afrotheria (e.g., elephant and aardvark), Xenarthra (e.g., armadillo and sloth), Euarchontoglires (e.g., primates and rodents), and Laurasiatheria (e.g., carnivores and ruminants), while also consistently supporting the monophyly of the latter two groups into a clade known as Boreoeutheria (for review, see Springer et al. 2004). Each of the four superordinal clades and Boreoeutheria have been verified by the identification of numerous rare genomic changes such as coding indels and retroposon insertions (Poux et al. 2002; de Jong et al. 2003; Nikaido et al. 2003; Thomas et al. 2003; Nishihara et al. 2005, 2006; Kriegs et al. 2006). This increasingly resolved phylogeny is being used as the primary guide for selecting new target mammalian genomes, and for promoting low coverage (2x) genomes to be sequenced at higher coverage, in an effort to identify the functional elements in the human genome (Margulies et al. 2005).

The exact placement of the root of the placental mammal tree has implications for inferring ancestral genomic sequences and genomes (Blanchette et al. 2004; Ma et al. 2006), as well as the early biogeographic history of placental mammals (Eizirik et al. 2001; Madsen et al. 2001; Murphy et al. 2001b; Springer et al. 2003; Hunter and Janis 2006). However, it has proved to be elusive (Murphy et al. 2004; Springer et al. 2004; Kriegs et al. 2006), and three hypotheses remain (Fig. 1): (1) a basal position for Afrotheria, termed Exafroplacentalia (Waddell et al. 2001); (2) a basal position for Xenarthra, called Epitheria (Shoshani and McKenna 1998); and (3) a monophyletic clade containing Afrotheria and Xenartha (Atlantogenata *sensu*) (Waddell et al. 1999) as a sister group to Boreoeutheria. While most molecular studies of large data sets have favored the Exafroplacentalia hypothesis (Madsen et al. 2001; Murphy et al. 2001a,b; Waddell et al. 2001; Amrine-Madsen et al. 2003; Waddell and Shelley 2003), analyses of subsets of genes, amino acid sequences, and more extensive taxon sampling within relevant clades have also supported Atlantogenata (Madsen et al. 2001; Murphy et al. 2001a; Delsuc et al. 2002; Douady and Douzery 2003; Waddell and Shelley 2003; Kjer and Honeycutt 2007). However, despite high Bayesian posterior probabilities for these different hypotheses, neither have been well supported with other measures, such as maximum likelihood bootstrap values, leaving the root of the placental tree unresolved (Murphy et al. 2004; Springer et al. 2004).

Rare genomic changes provide an independent assessment of hypotheses based on phylogenetic tree-building algorithms (Rokas and Holland 2000; Waddell et al. 2001; Shedlock et al. 2004; Bashir et al. 2005; Rokas and Carroll 2006). A recent study identified two L1MB5 retroelements that supported the Epitheria hypothesis by scanning available genome alignments (Kriegs et al. 2006), though this result was not considered statistically significant using the criterion developed for indels by Waddell et al.

[5]Corresponding author.
E-mail wmurphy@cvm.tamu.edu; fax (979) 845-9972.
Article published online before print. Article and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.5918807.

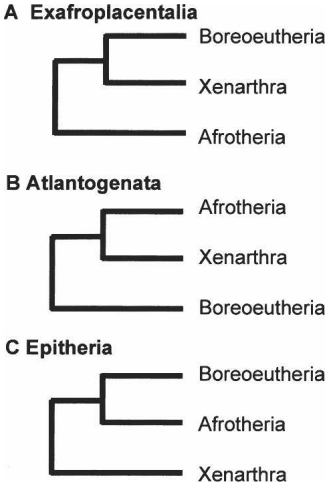

Figure 1. Three hypotheses for the basal relationships among placental mammal superordinal clades. The name for each hypothesis refers to the ingroup taxa.

(2001). One limitation of the L1MB5 elements reported by Kriegs et al. (2006) is that the alignments of the flanking regions are weakly conserved, making it difficult to rule out the possibility that the L1MB5 insertion predated theria and is now absent in xenarthrans due to a lineage-specific degradation of the retroelement. Though retroelements are presumably immune to such events, recent studies show that homoplasy can also occur in the form of targeted insertion, or lineage sorting of ancestral polymorphisms during rapid cladogenesis (Cantrell et al. 2001; Pecon-Slattery et al. 2004; Ludwig et al. 2005; van de Lagemaat et al. 2005; Yu and Zhang 2005; Nishihara et al. 2006). Therefore, the two elements supporting Epitheria found by Kriegs et al. (2006) require further validation, which is now facilitated by the recent completion of armadillo and elephant 2x genome sequences.

To resolve these issues we scanned genomic sequence assemblies of the African elephant (*Loxodonta africana*), nine-banded armadillo (*Dasypus novemcinctus*), and opossum (*Monodelphis domestica*) genomes, aligned to human RefSeq protein-coding exonic regions. Protein-coding alignments have the advantage of being more reliable for establishing sequence alignment orthology than noncoding alignments, particularly in early branches of the placental mammal tree. Coding regions are easily aligned to outgroup sequences to determine polarity of changes, and have been useful for identifying indels informative for early mammalian phylogeny (Poux et al. 2002; de Jong et al. 2003; Murphy et al. 2004; Springer et al. 2004; van Rheede et al. 2006). Furthermore, these characters may have selective advantages that promote rapid selective sweeps over shorter timescales and therefore may be less prone to lineage sorting than retroelements. Our analysis of >180,000 coding exons in multiple mammals supports the Atlantogenata hypothesis in which afrotherians and xenarthrans form a monophyletic group. This result is different than most of the largest molecular sequence-based phylogenetic trees, and the retroelement insertion analysis of Kriegs et al. (2006).

Results and Discussion

Coding indels support the Atlantogenata hypothesis

To computationally screen for informative exons, an alignment of the human genome with 16 other vertebrate genomes was downloaded from the UCSC Genome Browser (Kent et al. 2002). We identified 180,258 distinct human RefSeq protein-coding exons and extracted alignments of them to each of armadillo, elephant, and opossum from the 17-way alignment. Alignments were not allowed to include unsequenced regions (i.e., sequence data containing Ns). We found 96,612 human exons that align without a gap to armadillo, 104,332 for elephant, and 122,959 for opossum. Also, we found 5648 human exons that align to armadillo with a single internal gap of length divisible by three, 5723 for elephant, and 10,504 for opossum. We were particularly interested in human exons that align without a gap to one of armadillo, elephant, or opossum and have a single (internal, length divisible by three) gap at precisely the same position with respect to the other two species. We found 30 cases where human aligns to armadillo without a gap (i.e., Exafroplacentalia), while for human to elephant (i.e., Epitheria) and human to opossum (i.e., Atlantogenata) we found 42 and 55 cases, respectively. To take just one example, an instance of the third case (i.e., an exon where human and opossum align without a gap, but human aligns to armadillo and elephant with a gap at the same position) can be explained by a single insertion/deletion only under the Atlantogenata hypothesis (Fig. 1B); each of the other two scenarios for the eutherian root implies homoplasy.

After further electronic screening of these three lists for cases of paralogous alignments from the low-coverage armadillo and elephant assemblies (which were quite frequent), and ruling out cases of homoplasy by adding additional draft genome assemblies (i.e., mouse, rat, dog, and cow) and sequences from the trace archives (See Methods; Supplemental Table 1), the only remaining four candidates supported the Atlantogenata hypothesis. These were all verified by designing PCR primers surrounding the indel and sequencing in additional placental mammal taxa. Sequences for the best taxon-represented case, *PTPRB* exon 9, were recovered from 48 amniotes, all of which supported interpretation of the indel as a unique event in the common ancestor of Afrotheria and Xenarthra (Fig. 2). Similar compilations of sequence traces and PCR-sequencing verification for the other three candidate exons, *ZNF367* exon 5, *C14orf121* exon 35, and *LAMC2* exon 13, confirmed the presence of a single amino acid insertion or deletion event in all afrotherians and xenarthrans tested that was absent in all boreoeutherians and outgroup species (Figs. 3, 4). For two of the indels, *ZNF367* and *LAMC2*, subsequent loss or gain of the indel in other terminal lineages was noted (Figs. 3, 4). However, our broad taxon sampling suggests that these recurrences are extremely rare. In contrast, we found no indels to support the Exafricoplacentalia or Epitheria hypotheses.

To corroborate our findings from the exons, we examined the 17-way multispecies alignment at the UCSC Genome Browser for potential retroelement insertions favoring the three hypotheses, also requiring good flanking alignments in the opossum assembly to rule out lineage-specific degradation (See Methods). We found two insertions that supported Atlantogenata and none that supported Epitheria or Exafricomammalia under similar search criteria. The first example is a L1MB5 insertion in the *OTC* gene on chromosome X (Fig. 5A). This element is bounded by a conserved coding exon on the left and a human lod (conserved noncoding region as defined in the UCSC browser) on the right. This region aligns very well to the opossum assembly, which is contiguous to non-insert boreoeutherian species. Trace archive sequences and PCR analysis of 11 additional taxa confirm that this L1MB5 insertion is shared by all extant afrotherian and xenarthran lineages tested (Fig. 5A).

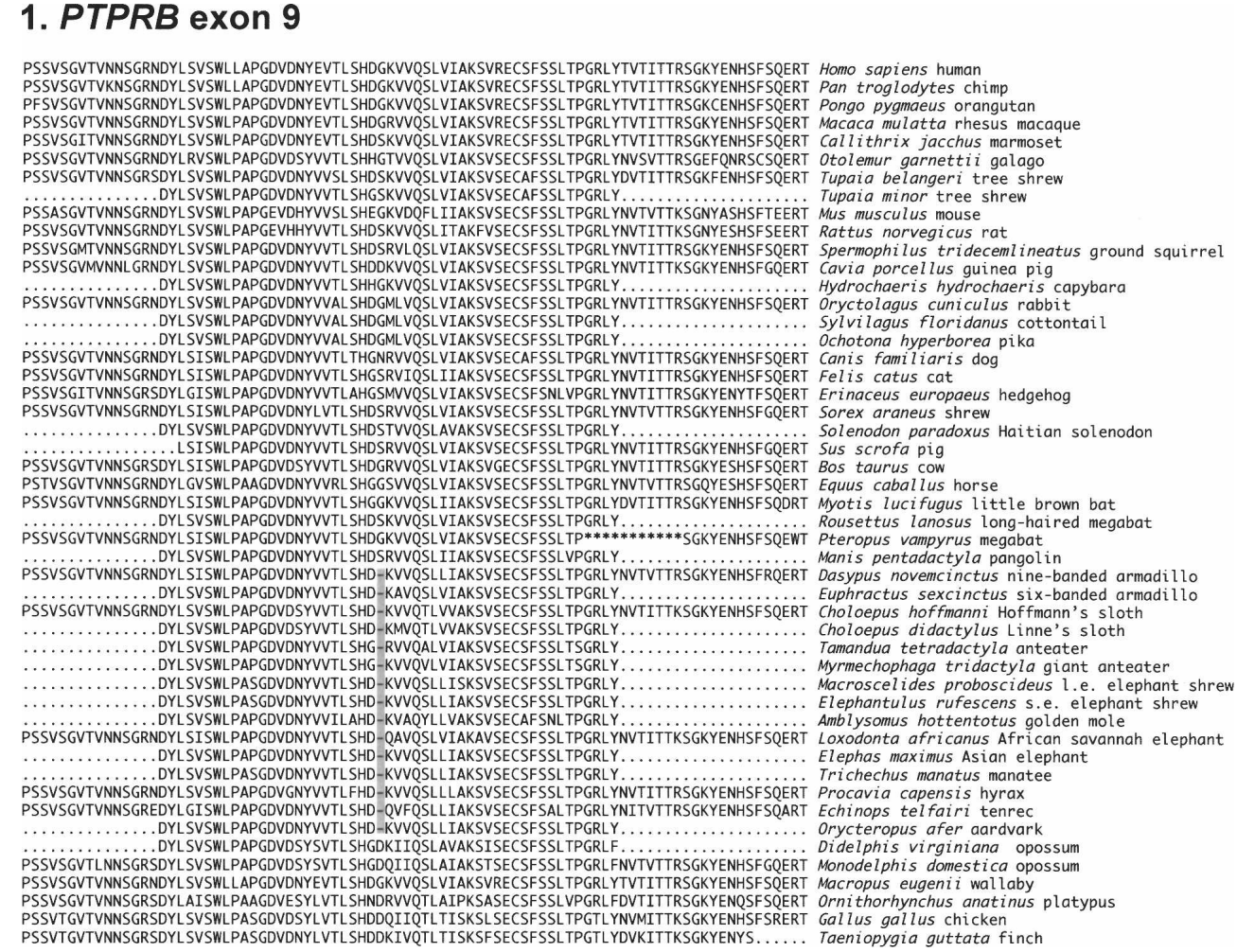

Figure 2. Alignment of *PTPRB* exon 9 sequences from 49 vertebrate species. A single amino acid deletion shared by all xenarthrans and afrotherians is highlighted with a box. Full-length amino acid sequences are derived from draft genome sequences or traces. Sequences with truncated ends were generated by PCR using conserved primers in the exon flanking the indel. Asterisks indicate gaps in the sequence assemblies or traces.

The second L1MB5 insertion is within a large intron of a coding gene, *RABGAP1L*, on human chromosome 1 (Fig. 5B). Trace archive and PCR-generated sequences from four afrotherians and three xenarthrans shared a L1MB5 insertion, which was absent in all 16 boreoeutherian species examined. The opossum alignment showed no evidence for such an insertion. Therefore, together with the four coding indels, these two retroposon insertions provide strong support for the Atlantogenata hypothesis.

A revised molecular timescale for placental evolution

Given the support for Atlantogenata, we re-estimated a molecular timescale for placental mammals using a large nuclear+mitochondrial data set (Springer et al. 2003; Roca et al. 2004) and a Bayesian relaxed clock approach (Thorne et al. 1998; Kishino et al. 2001), adding a constraint on the monophyly of Afrotheria+Xenartha (Atlantogenata). We also allowed the tree to support the recent Pegasoferae hypothesis (Nishihara et al. 2006), which places bats in a monophyletic group with carnivores, pangolins, and perissodactyls. Under this topology we estimate the split between Afrotheria and Xenarthra occurred 103 million years ago (Mya) (95% credibility interval, CI = 95 – 114 Mya), just two million years after the point estimate for the radiation of crown placental mammals 105 Mya (95% CI = 96 – 115 Mya) (Fig. 6). Other dated splits include 97 Mya (95% CI = 90 – 106 Mya) for Boreoeutheria, 91 Mya (95% CI = 84 – 99 Mya) for Euarchontoglires, and 87 Mya (95% CI = 82 – 93 Mya) for Laurasiatheria. The remaining ordinal divergence times (Supplemental Table 2) are similar to those found in previous studies (Douady and Douzery 2003; Springer et al. 2003; Roca et al. 2004).

The Atlantogenata hypothesis is consistent with a strict vicariant scenario where the rifting of South America and Africa ~100–120 Mya (Smith et al. 1994; Hay et al. 1999) formed the ancestors of Xenarthra and Afrotheria, respectively. Earlier phylogenetic hypotheses with either Afrotheria or Xenarthra in basal positions implied a southern origin for placental mammals (Eizirik et al. 2001; Madsen et al. 2001; Murphy et al. 2001b). The new phylogeny with reciprocal monophyly of northern (Boreoeutheria) and southern (Atlantogenata) clades, however, is not at odds with a northern origin for crown placental mammals, followed by dispersal to the south (Hunter and Janis 2006).

Homoplasy confusing the basal placental mammal issue always remains a possibility; however, the time available for sepa-

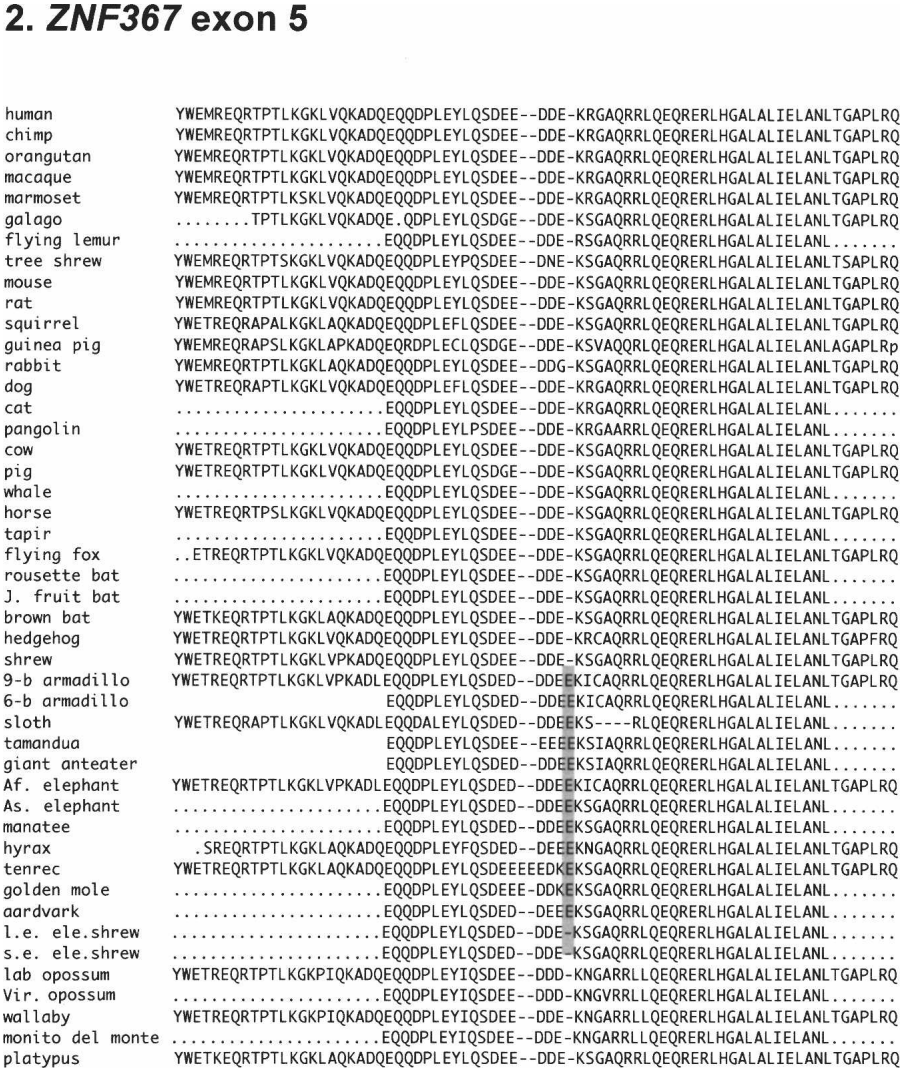

Figure 3. Alignment of *ZNF367* exon 5 sequences from 46 mammals. A single amino acid insertion shared by all xenarthrans and afrotherians is highlighted with a box. Full-length amino acid sequences are derived from full genome sequences or traces. Sequences with truncated ends were generated by PCR using conserved primers in the exon flanking the indel.

rate events in both stem afrotheres and stem xenarthrans is a small fraction of this time of observed nonrecurrence (25 + 29 = 54 million years, or 3% of 1.5 billion years of sampled placental evolutionary time). It is worth noting that the proposed phylogeny with approximately two million years of common stem shared by Afrotheria and Xenarthra is an intrinsically unfavorable situation for recurrent mutation, being only 4% of the unobservable stem lengths (lack of extant species) and 1% of the overall sampled placental timescale. The short stem branch for Atlantogenata likely explains why previous molecular studies have had difficulty recovering robust support for the placental root. The closest outgroup taxon to placental mammals, marsupials, is still quite distant with roughly 185–225 million years of divergence on stem branches between crown placentals and crown marsupials (Kumar and Hedges 1998; Woodburne et al. 2003; van Rheede et al. 2006) which would favor rooting on the longer branch to Afrotheria or Xenarthra. It also explains why two previous studies attempting to identify retroposon insertions diagnostic for placental superordinal clades (Kriegs et al. 2006; Nishihara et al. 2006) found no evidence for the Atlantogenata hypothesis.

Evaluation of Kriegs et al.'s two insertions supporting Epitheria

Retroelements have strong potential to resolve difficult nodes of a phylogeny (Shedlock and Okada 2000; Murphy et al. 2004; Springer et al. 2004; Bashir et al. 2005; Kriegs et al. 2006; Nishihara et al. 2006; Rokas and Carroll 2006). Their advantages are immunity to later identical events and probably a higher event (insertion) rate than coding indels. Initially these events were considered nearly perfect characters; however, various validation issues arise in firmly establishing orthology, addressing deletional homoplasy, ruling out quasi hotspots of independent insertion of similar length fragments from a given master element, proving absence in the immediate outgroup, and determining status in a wide range of relevant species.

RepeatMasker may very well report precise fragment end points but it cannot determine them to statistical significance in a gapped alignment situation. Indeed, the two L1MB5 examples reported by Kriegs et al. (2006) to support Epitheria exhibit a distinctly ragged set of end points when tracked across various species. In our search, both of Kriegs L1MB5s were excluded because they lack a clean insertion in armadillo (i.e., the boundaries of the gap in armadillo are offset relative to the boundaries of the annotated repeat in the multispecies alignment). The second repeat reported by Kriegs in support of Epitheria (their 2b) resides in a region of human chromosome 15 that does show a good alignment to the opossum genome. Though probably too divergent to be reliably recognized by RepeatMasker, the presence of opossum sequence spanning the region covered by the L1MB5 in all nonxenarthran eutherians suggests lineage-specific loss as a probable explanation for the gap seen in xenarthrans, rather than support for the Epitheria hypothesis (Supplemental Fig. 1).

While small lineage-specific indels can be invoked to explain ragged end points, this undercuts the notion that common end points of a retroposon fragment are diagnostic of orthology. Further, if an L1MB5 element is observed to be peppered with small indels, why not the occasional large deletion? This does not require a mechanism for precise excision (e.g., nonhomologous recombination with a near-contemporaneous same-master similar fragment L1MB5 elsewhere in the genome) because RepeatMasker cannot find residual fragments below a certain length (~25 bp). Should this take out enough of the repeat to make it operationally unrecognizable to RepeatMasker, and occur in a common ancestor of a superordinal clade, then the presence/

3. *C14orf121* exon 35

```
Human                GTEGSEPGEGGPAPGTAQQP-RVHGVALPGLERAKGWSFDGKRE
chimp                GTEGSEPGEGGPAPGTAQQP-RVHGVALPGLERAKGWSFDGKRE
orangutan            GTEGSEPGEGGPAPGTAQQP-RVHGVALPGLERAKGWSFDGKRE
macaque              GTEVSEPGEGGPAPGIAQQP-RVHGVALPGLERAKGWSFDGKRE
marmoset             GTEGSEPGEDGPAPGTAQQP-RVHGVALPGLERAKGWSFDGKRE
bushbaby             GTEGPEPGEGGPAPGPAQQS-RAHGVSLPGLERAKGWSFDGKRE
tree shrew           GTEGSEPGEGDPTPTTTQLP-RVHGVALPGLGRAKGWSFDGKQE
mouse                GTERLEAGEGAPAPGTAQQP-RVHGVALPGLGRTKGWSFDGKRE
rat                  GTERLEAGDGAPAPGTAQPP-RVHGVALPGLGRTKGWSFDGKRE
ground squirrel      GTEGSEPGEGDPVPGTAQQP-RAHGVALPGLGRAKGWSFDGKRE
guinea pig           GTEGSEAGDGGTAPGMAQQP-AVHGVALPGLGRAKGWSFDGKQE
beaver               GTEGAEPGEGGPAPGTAQQP-RIHG..................
llama                GTEGAEPGEGGQAPGTAQQP-RVHG..................
flying lemur         ..EGSEPGEGSPAPGTAQQP-RVHA..................
hedgehog             GTEGAEPRDGGPAPGMTQQP-RVHGVALPELGRAKGWSFDGKQR
horse                GTEGTEPGEGVQAPGTAQQP-RVHGVALPGLGRAKGWSFDGKRE
cow                  GTEGSEPGEGGPAPGTAQQP-RVHGVALPGLERAKGWSFDGKRE
pig                  GTEGAEPGEGGQAPATAQQP-RVHGVALPGLGRAKGWSFDGKRE
little brown bat     GSEGSEPGEGAQAPGTAQQP-RVQGVALPGLGRTKGWSFDGKQE
flying fox           GTEGSEPGDGGQTPVTAQQP-RVHGVALPGLGRAKGWSFDGKQE
dog                  GTEGAEPGEGGLAPGTAQQP-RVHGVALPGLGRAKGWSFDGKRE
cat                  GTEGAEPGEGGLAPGTAQQP-RVHGVGPSGVGKSQG........
9-banded armadillo   GMEGSEPGEGGPAHGTTQQPTRVHSIALPGLGRAKGWSFDGKRE
Hoffmann's sloth     GTDGSEPGEGGPAQGTTQQPTRVHGVALSGLGRAKGWSFDGKRE
Linne's sloth        GTEGAEPGEGGPAQGTTQQPTRVHG...................
giant anteater       CTEGAEPGEGGPAHGTTQQPARVHG...................
tamandua             GTEGSEPGEGGPAHGPTQQPARVHG...................
African elephant     GTEVSEPGEGGPALGALQQSTRVHGVALPGLGRAKGWSFDGQRE
manatee              .TEGAEPGEGGLAPGTSQQSTRVHGVA.................
hyrax                GTEVSEPGEGVPALGASQQPARVHGVALPGLGRAKGWSFDGKQE
tenrec               GTEVSEPGEGVPALGASQQPNRVHAVALPGLGRAKGWSFDGKRE
opossum              GMEGAEPGEGVPVSAVMQSA-RVHGIALPGMGRAKGWSFDGKQE
monito del monte     .TEGAEPGEGVPVSGVMQPT-RVHGIA.................
platypus             -------SSSGPEAALLQPP-RLQGIALPGLGRVKGWTFEGRRE
```

4. *LAMC2* exon 13

```
human                MDQFMQQLQRMEALISKAQGGDGVVPDTELEGRMQQAEQALQDILRDAQISE
chimp                MDQFMQQLQRMEALISKAQGGDGVVPDTELEGRMQQAEQALQDILRDAQISE
orangutan            MDQFMQQLQRMEALISKAQGGDGAVPDTELEGRMQQAEQALQDILRDAQISE
macaque              MDQFMQQLQRMEALISKAQGGDGAVLDTELEGRMQQAEQALRDILRDAQISE
bushbaby             MDQFMQQLQSLEALISKAQSGGGAVPNTELEDRMQQAEQALRDILREAQISE
mouse                MDQFTQQLQSLEALVSKAQGGGGTVPSGELEGRMQQAEQALRDILREAQISE
rat                  MDQFAQQLQNLEALVSKAQGGNGAVPSEELEGRMQQAEQALRDILREAQISE
ground squirrel      MDQFMQQLQSLEALISKAQGGGAAVPGSELEGGMQQAEQALRDILREAQISE
dolphin              MDQFMQQLQSLEALVSKAQGGGGAVPDAELEGRMQQAEQALQDILREAQISE
dog                  MDQFMQQLQSLEALVSRVQGGGGAVPDAELEGRMQQAEQALQELLREAQISE
cat                  MDQFRQQLQSLEALVSRAQSGGGAVPSTELEGRMRQAEQTLQDILREAQISE
hedgehog             VDLFMQQLQRLEALVSRAQGAGGAVSNSELEGRMQQAEQALQDILREAQISE
solenodon            .............LISQAQGSGGTLPNTELEGRMQ..................
horse                MDQFMQQLQILEALISKAQGG--AVPNAELEGRMQQAEQALRDILREAQISQ
cow                  MDQFMQQLESLETLLSKAQAGGGAVPDTELEGRMQQAEQALQDLLREAQISE
pig                  MDQFMQQLQSLETLISKAQDGGGTIPNTELESRMQQTEQTLEDILREAQIAE
llama                .............LILMAQGGGGAVPSTELEGRMQ..................
little brown bat     MGQFVQQLQSLEDLISKAQGGGGGAPNAELEGRMQQAEQALRDMLRDAQISQ
9-banded armadillo   MDQFTQQLQSLEALVSEARDG-GAVPNAELEGRIQQAEEALWNILREAQISE
sloth                MDQFMQQLQSLEALISEVQGG-GGVAHPELEGKMQQAEQALQDILREAQISE
African elephant     MDQFMQQLQSLEALISKAQEG-GGVPHPELEGKMQQVEQALQDILREAQISE
manatee              .............MVSNTQDG-GAVPNAELEGRMQ..................
hyrax                MDQFTQQLQSLEALVSKAQEG-GAVPHAELEVRTQEAEEALRDILREAQVSE
elephant shrew       .............LVSKAQGG-GAIPSAELEGRMQ.................
tenrec               MDQFKHQLQALEALMAKAQGG-GAVPNSELEGRMQQAEEALQDVLRDAQVSE
opossum              LEQYLQQLQSVEALVLKVQAGGGSLPNAELEGRLKEAERTLQDILREAQISE
wallaby              LEQYLQQLQSLEALVLKVQAGGGSLPNAELEGRLKEAEGTLQDILREAQISE
platypus             MDQLWLQIRDLEVHL-----GGAPGPNTEVERRMHQVEETLQNILREAQISE
```

Figure 4. Alignment of *C14orf121* exon 35 sequences from 34 mammals (*upper* panel), and *LAMC2* exon 13 sequences from 28 mammals (*lower* panel). The indels shared by xenarthrans and afrotherians are highlighted with boxes. Full-length amino acid sequences are derived from draft genome sequences or traces. Sequences with truncated ends were generated by PCR using conserved primers in the exon flanking the indel.

absence of the repeat element character will be misread, leading to incorrect phylogenetic inference. In the case at hand, xenarthrans have nearly 30 million years of stem prior to first divergence of extant species, which may provide ample time for the two insertions observed by Kriegs et al. (2006) to have degraded beyond recognition.

An additional explanation for the two insertion examples reported by Kriegs et al. (2006) is incomplete lineage sorting of an ancestral polymorphic insert during a very short divergence time (Shedlock et al. 2004). The approximately two million year interval (based on point estimates) between the ancestor of placental mammals and the common ancestor of Atlantogenata shown in our mammalian divergence time analysis (Fig. 6) provides just such an example. A similar situation was reported by Nishihara et al. (2006) for the Pegasoferae clade. In that case one discrepant L1MB5 insertion was found that contradicted four other insertions supporting monophyly of bats, perissodactyls, and carnivores. Our molecular timescale under the Pegasoferae hypothesis shows that these three lineages also diverged during a remarkably rapid time frame: The point estimates for these branches are less than two million years apart. Under such conditions of rapid divergence events, it would not be uncommon to find cases where fixation of the original insertion through lineage sorting is homoplastic (Shedlock et al. 2004; Nishihara et al. 2005).

Other pivotal assumptions of retroposon insert analysis, such as random insertion of L1 repeats in a large genome having negligible odds of recurring (nearby or same-spot insertion, homoplasy), conflict with two recent observations. First, two human L1s from the same active source have been observed to insert within 87 bp in unrelated patients (Ludwig et al. 2005). Second, L1 insertions have distinct genomic biases whose details are imperfectly known (Graham and Boissinot 2006). This has serious homoplasy implications for the use of this repeat class in inferring ancestral topologies because it would be difficult to distinguish single from multiple scenarios after 100 million years of subsequent mutation. While retroposons clearly are a powerful resource for phylogenomics, careful PCR-based validation, followed by interpretation in light of molecular divergence estimates, is necessary to fully test the reliability of such data mined from genomic databases. Though cases of possible lineage sorting of ancestral polymorphism or other types of homoplasy will be less common or even negligible on longer branches (such as those leading to Euarchontoglires or Laurasiatheria), short nodes inside of rapidly diversifying clades like Laurasiatheria and Afrotheria (Nishihara et al. 2005, 2006) must be thoroughly tested by other types of data such as exonic indels.

Conclusions

The recent studies by Kriegs et al. (2006) and Nishihara et al. (2006) firmly establish the monophyly of Laurasiatheria, Euarchontoglires, Boreoeutheria, and Afrotheria, confirming results based on phylogenetic analysis of most large nuclear and mitochondrial gene data sets. Our present analysis of coding indels and retroposon insertions provides strong evidence for the root of the placental tree between Atlantogenata and Boreoeutheria, a finding that accords well with the vicariant separation of Africa and South America in the Cretaceous. The current influx of genome sequence data from nearly all orders of placental mammals should soon permit the complete unraveling of the phylogenetic relationships among the early interordinal branches of the mammalian tree.

Given the incomplete nature of the low-coverage elephant and armadillo assemblies and the need for coverage at the same regions across genomes, it remains a possibility that the sample we examined was unrepresentative or somehow biased in a manner that we could not control for. This issue will be worth revisiting when more complete draft assemblies from these and addi-

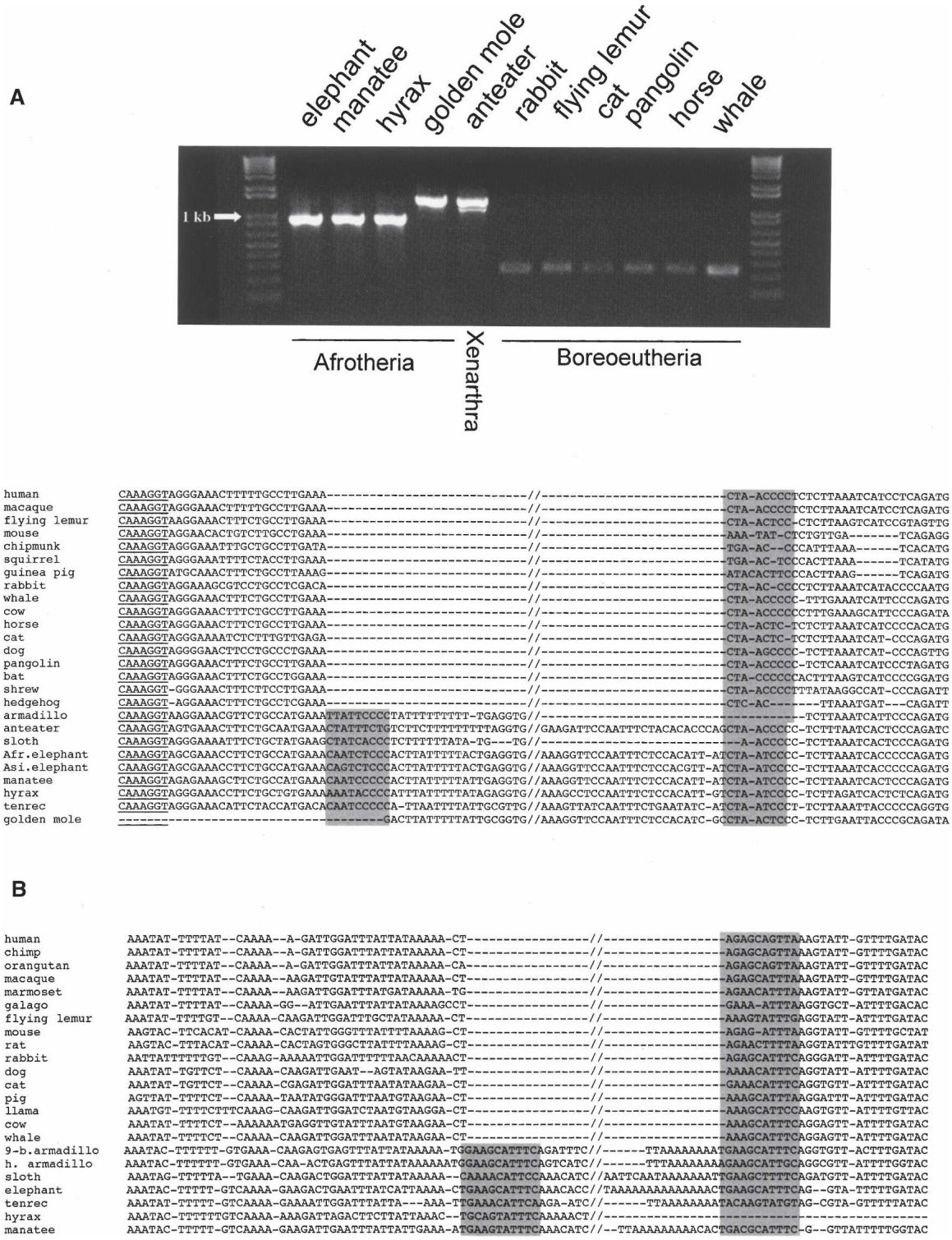

Figure 5. (*A*) L1MB5 insertion in an intron of the *OTC* gene. The *upper* gel image shows PCR products obtained by amplifying different mammalian genomic DNAs with conserved primers flanking the retroposon insertion site. *Below* is a partial alignment of the exonic (underlined)/intronic sequences flanking the L1MB5 insertion from 26 placental mammals. The central portion of the inserted sequence has been edited, as shown by the double slash marks. (*B*) A multispecies alignment showing the second L1MB5 insertion in an intron of the *RABGAP1L* gene from 23 placental mammals. Direct repeats are highlighted in gray.

tional species become available. These data will facilitate the assembly of larger data sets for analysis and will increase the ease with which corroborating rare genomic changes are identified (Murphy et al. 2004; Kriegs et al. 2006; Nishihara et al. 2006; Rokas and Carroll 2006). Moreover, the data will allow an accurate prediction of the evolutionary history of those species at all

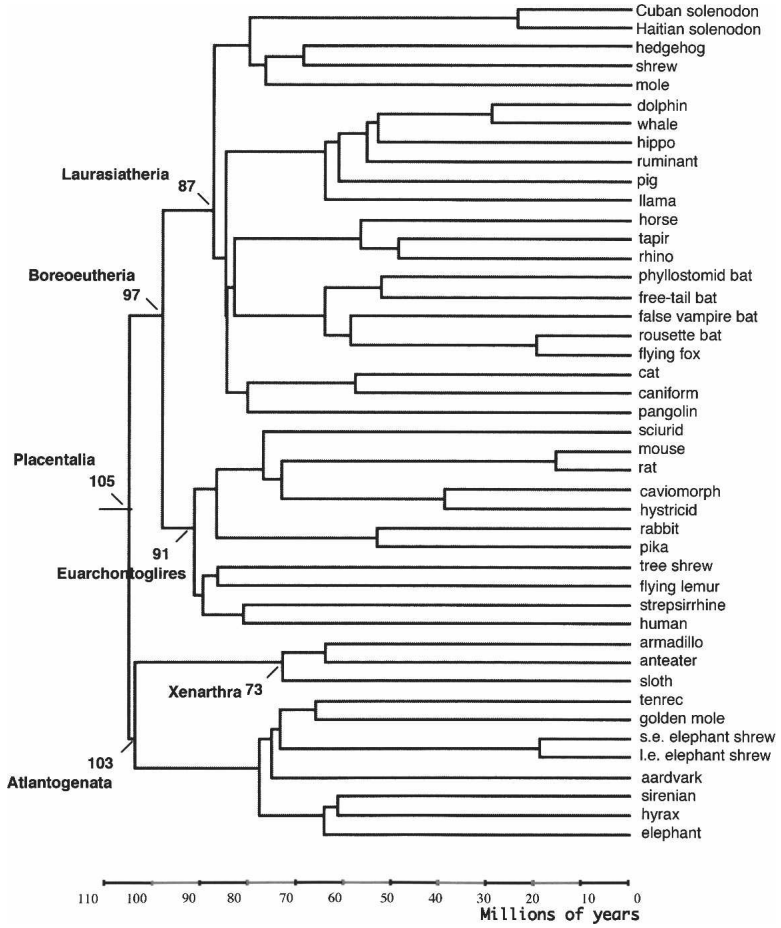

Figure 6. A molecular timescale for placental mammals based on the data set from Roca et al. (2004), 13 fossil constraints (Springer et al. 2003), and a mean prior of 105 Mya for the placental root. Divergence estimates are shown for several key superordinal clades (for a full list of divergence times and confidence intervals, see Supplemental Table 2).

scales of evolutionary events, down to nucleotide resolution (Blanchette et al. 2004). At one level, phylogenetic relationships correspond to the largest-scale evolutionary event, i.e., speciation, in a long list that includes chromosome fission and fusion, large-scale duplication, insertion of interspersed repeats, deletions, expansion/contraction of tandem repeats, and nucleotide substitutions, among others. However, correct determination of phylogenetic relationships has a special importance for the goal of thoroughly understanding mammalian genome evolution, first in informing decisions about which sequence data should be gathered, and later to determine the order that computational procedures are applied to the data.

Methods

Identification of potential coding indels in human–armadillo–elephant–opossum whole-genome alignments

The alignment of 17 vertebrate genomes was downloaded from http://genome.ucsc.edu/ in April 2006. It included sequence data from human assembly hg18 (March 2006), armadillo assembly dasNov1 (May 2005), elephant assembly loxAfr1 (May 2005), and opossum assembly monDom4 (January 2006) (E. Lander, pers. comm.). A special-purpose program was written to identify which protein-coding intervals align between human and each of armadillo, elephant, and opossum, with at most one gap, and to compare the locations of those gaps, as described in Results. Source code for that program is freely available as part of the package called "Phylogenomic Tools" (Rosenbloom et al., in press), which is available at http://www.bx.psu.edu/miller_lab/.

Identification of diagnostic retroelement insertions

For the Atlantogenata hypothesis we scanned the 17-way alignment for locations where armadillo and elephant had an extra segment of between 150 and 2000 additional nucleotides in a region that aligned contiguously in human, dog, and mouse; the region flanking the repeat was required to have a contiguous alignment in opossum to rule out lineage-specific degradation. Sequences spanning that region in human, mouse, dog, armadillo, elephant, and opossum were aligned with BLASTZ (Schwartz et al. 2003), and the alignments interactively inspected using the laj program (Wilson et al. 2001), both of which are freely available at http://www.bx.psu.edu/miller_lab/. RepeatMasker was run on the elephant sequence, and the putative inserted region was required to coincide closely with an annotated repeat element whose family and level of divergence from the consensus sequence are consistent with having been inserted around 100 Mya. Full sequences for each retroelement and their flanks can be found in Supplemental Figures 2 and 3.

A similar procedure was run for the Epitheria and Exafricoplacentalia hypotheses. We scanned the 17-way alignment for a non-Alu interspersed repeat of length between 150 and 2000 bp found in human, dog, mouse, and armadillo that does not align to elephant (in the case of the Exafricoplacentalia hypothesis) or found in human, dog, mouse, and elephant that does not align to armadillo (in the case of the Epitheria hypothesis). The putative inserted region was required to coincide closely with an annotated repeat element whose family and level of divergence from the consensus sequence are consistent with having been inserted around 100 Mya. The region flanking the repeat was required to have a contiguous alignment in opossum to rule out lineage-specific degradation.

In silico screening/validation of indels with additional mammalian genomes

Candidates identified in the initial comparative genomics screen still include cases of cross-matched paralogs (arising from incomplete elephant and armadillo assemblies), cases of indels within simple repetitive sequence (conducive to multiple homoplastic events through replication slippage), situations with highly diverged sequence (indicating little constraint on indels), and ex-

amples exhibiting inconsistent phylogeny within the UCSC genome browser species (e.g., indels also present in species like dog or cow that were not part of the original screening).

A second round of rapid screening was thus implemented to reduce these sources of artifacts. DNA from the four species representing the putative orthologous exon was required to have comparable scores upon BLAT against the human genome at the UCSC browser (no worse than 50% of human vs. human), no paralogous chromosomal location with higher scores, and consistent reciprocal best-BLAT. The comparative genomics and xeno mRNA tracks were then revisited to verify phylogenetic consistency in species available as of August 2006 (cow, chimp, macaque, mouse, rat, rabbit, guinea pig, tenrec, opossum, platypus, chicken, and frog).

We further required two-sided synteny of flanking genes in species where contig availability made this feasible. While not an infallible guide to orthology even when coupled with best reciprocal BLAT hit plus splice site location and phase (because of segmental duplications and incomplete genomes), this does support reliable tracking of the indel back into species in which the exon sequence has significantly diverged. Surviving exon candidates were screened about the indel site for anomalous composition, direct repeats, palindromes, stem–loop potential, and RNA secondary structure that might favor recurring events (homoplasy) using the appropriate UCSC browser tracks and DotPlot (http://entelechon.de/bionautics/dotplot.php3).

Next, a concerted effort was made—even for those supporting other topologies—to confirm the indel in the greatest possible set of additional species, with special emphasis on the immediate outgroup (marsupials), the deepest nodes of afrotheres and xenarthrans, dispersed representatives from all 18 mammalian orders, particularly those not already represented in the trace archives, and obtaining a maximal amount of elapsed branch length evolutionary time without indel recurrence. None of the candidates extended reliably past tetrapods because of varying divergence and differentially expanded gene families; platypus works overall as a cutoff to relevance, though some candidates could be extended to chicken, zebrafinch, anole, and frog.

Promising indels were broadened from the core species of the screening set to the widest possible set of orthologs available at NCBI databases, notably trace archives, wgs (traces assembled into small contigs), est_others, nr, htgs, and draft assemblies at the UCSC genome browser. Although more prone to frameshifts and base pair errors than PCR, trace indels are commonly a single base pair in length within compositionally simple regions, thus causing minimal confusion to the 3n base pair coding indels considered here (which are not in simple repeat regions). In most species, multiple traces are available with high-quality interiors completely spanning the short candidate exons and validating GT-AG splice sites and phasing (important in avoiding processed pseudogenes). Less favorable situations have partial coverage from a single trace or two traces needing tiling to span the exon. It is not possible to establish synteny from traces in advance of draft genome assembly. These species were then supplemented by PCR as needed to better sample afrotherians, xenarthrans, and placental orders underrepresented to date in genomic sequencing projects.

PCR validation and sequencing of candidate indels

For those indels that were supported by additional in silico alignment to whole-genome assemblies and trace archive sequences, we attempted to validate each using PCR-based sequencing of orthologous regions in additional species. Conserved PCR primers were designed in stretches of exonic sequence flanking the candidate indels, though occasionally primers were extended into conserved flanking intronic sequence (Supplemental Table 3). PCR was tested on a panel of genomic DNAs from species spanning all four eutherian superordinal groups, using a touchdown PCR protocol (annealing temperature from 60°C to 50°C) and varying magnesium chloride concentration from 1.25 to 2.5 mM. For those species which produced a clean band following agarose gel electrophoresis, we directly sequenced each product after clean-up with Microcon PCR purification (Millipore) or ExoI/shrimp alkaline phosphatase digestion. Sequencing reactions were performed with Big-Dye Terminator ready reaction mix and resolved on an ABI-3730 sequencer (Applied Biosystems). DNA sequences were trimmed and aligned in Sequencher (Gene Codes, Inc.). PCR sequences and genome assembly/trace fragments to which they align are available in Supplemental Material or in GenBank. ClustalW (Chenna et al. 2003), MALIGN (Wheeler 2003), TCoffee (Notredame et al. 2000), and DB BLAST alignments (http://www.proweb.org/proweb/Tools/WU-blast.html) on both DNA and protein were explored to see if they provided any insight or objectivity on gap placement beyond what hand-alignment could provide.

Molecular divergence time estimation

We computed a molecular timescale for placental mammal evolution using the large nuclear+mitochondrial gene data set from Roca et al. (2004), which is the same data set used by Springer et al. (2003) but includes the two solenodon species. We identified the best topology under the Atlantogenata hypothesis, while also assuming the monophyly of the recently discovered clade Pegasoferae (Nishihara et al. 2006). We used the Bayesian relaxed clock approach of Thorne and colleagues (Thorne et al. 1998; Kishino et al. 2001) as implemented in *estbranches* and *divtime5b*. Fossil constraints and run parameters were the same as those used in Springer et al. (2003) and Roca et al. (2004).

Acknowledgments

We thank the Broad Institute for making the armadillo, elephant, and opossum assemblies available prior to publication. We also thank Agencourt Biosciences, the Baylor Genome Sequencing Center, and the Washington University Genome Center for generating raw genomic sequence reads available in the trace archives, as well as Kate Rosenbloom, Brian Raney, and Jim Kent at UCSC for making the 17-way alignment available. We thank Alison Wilkerson for technical support and Jan Janecka for comments on an earlier version of the manuscript. This work was supported by funds from Texas A&M University, grants EF0629849 (W.J.M.) and EF0629860 (M.S.S.) from the National Science Foundation, and grant HG02238 from the National Human Genome Research Institute (W.M.).

References

Amrine-Madsen, H., Koepfli, K.P., Wayne, R.K., and Springer, M.S. 2003. A new phylogenetic marker, apolipoprotein B, provides compelling evidence for eutherian relationships. *Mol. Phylogenet. Evol.* **28:** 225–240.

Bashir, A., Ye, C., Price, A.L., and Bafna, V. 2005. Orthologous repeats and mammalian phylogenetic inference. *Genome Res.* **15:** 998–1006.

Blanchette, M., Green, E.D., Miller, W., and Haussler, D. 2004. Reconstructing large regions of an ancestral mammalian genome in silico. *Genome Res.* **14:** 2412–2423.

Cantrell, M.A., Filanoski, B.J., Ingermann, A.R., Olsson, K., DiLuglio, N., Lister, Z., and Wichman, H.A. 2001. An ancient retrovirus-like element contains hotspots for SINE insertion. *Genetics* **158:** 769–777.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. 2003. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **31:** 3497–3500.

de Jong, W.W., van Dijk, M.A.M., Poux, C., Kappe, G., van Rheede, T., and Madsen, O. 2003. Indels in protein-coding sequences of Euarchontoglires constrain the rooting of the eutherian tree. *Mol. Phylogenet. Evol.* **28:** 328–340.

Delsuc, F., Scally, M., Madsen, O., Stanhope, M.J., de Jong, W.W., Catzeflis, F.M., Springer, M.S., and Douzery, E.J. 2002. Molecular phylogeny of living xenarthrans and the impact of character and taxon sampling on the placental tree rooting. *Mol. Biol. Evol.* **19:** 1656–1671.

Douady, C.J. and Douzery, E.J.P. 2003. Molecular estimation of eulipotyphlan divergence times and the evolution of "Insectivora." *Mol. Phylogenet. Evol.* **28:** 285–296.

Eizirik, E., Murphy, W.J., and O'Brien, S.J. 2001. Molecular dating and biogeography of the early placental mammal radiation. *J. Hered.* **92:** 212–219.

Graham, T. and Boissinot, S. 2006. The genomic distribution of L1 elements: The role of insertion bias and natural selection. *J. Biomed. Biotechnol.* **1:** 75327.

Hay, W.W., DeConto, R.M., Wold, C.N., Wilson, K.M., Voigt, S., Schulz, M., Wold-Rossby, A., Dullo, W.-C., Ronov, A.B., Balukhovsky, A.N., et al. 1999. Alternative global Cretaceous paleogeography. In *Evolution of the Cretaceous ocean/climate system* (eds. E. Barrera and C. Johnson), pp. 1–48. The Geological Society of America, Boulder, CO.

Hunter, J.P. and Janis, C.M. 2006. "Garden of Eden" or "Fool's Paradise"? Phylogeny, dispersal, and the southern continent hypothesis of placental mammal origins. *Paleobiology* **32:** 339–344.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. 2002. The human genome browser at UCSC. *Genome Res.* **12:** 996–1006.

Kishino, H., Thorne, J.L., and Bruno, W.J. 2001. Performance of a divergence time estimation method under a probabilistic model of rate estimation. *Mol. Biol. Evol.* **18:** 352–361.

Kjer, K.M. and Honeycutt, R.L. 2007. Site specific rates of mitochondrial genomes and the phylogeny of eutheria. *BMC Evol. Biol* **7:** 8.

Kriegs, J.O., Churakov, G., Kiefmann, M., Jordan, U., Brosius, J., and Schmitz, J. 2006. Retroposed elements as archives for the evolutionary history of placental mammals. *PLoS Biol.* **4:** e91.

Kumar, S. and Hedges, S.B. 1998. A molecular timescale for vertebrate evolution. *Nature* **392:** 917–920.

Ludwig, A., Rozhdestvensky, T.S., Kuryshev, V.Y., Schmitz, J., and Brosius, J. 2005. An unusual primate locus that attracted two independent Alu insertions and facilitates their transcription. *J. Mol. Biol.* **350:** 200–214.

Ma, J., Zhang, L., Suh, B.B., Raney, B.J., Burhans, R.C., Kent, W.J., Blanchette, M., Haussler, D., and Miller, W. 2006. Reconstructing contiguous regions of an ancestral genome. *Genome Res.* **16:** 1557–1565.

Madsen, O., Scally, M., Douady, C.J., Kao, D.J., DeBry, R.W., Adkins, R., Amrine, H.M., Stanhope, M.J., de Jong, W.W., and Springer, M.S. 2001. Parallel adaptive radiations in two major clades of placental mammals. *Nature* **409:** 610–614.

Margulies, E.H., Vinson, J.P., NISC Comparative Sequencing Program, Miller, W., Jaffe, D.B., Lindblad-Toh, K., Chang, J.L., Green, E.D., Lander, E.S., Mullikin, J.C., et al. 2005. An initial strategy for the systematic identification of functional elements in the human genome by low-redundancy comparative sequencing. *Proc. Natl. Acad. Sci.* **102:** 4795–4800.

Murphy, W.J., Eizirik, E., Johnson, W.E., Zhang, Y.P., Ryder, O.A., and O'Brien, S.J. 2001a. Molecular phylogenetics and the origins of placental mammals. *Nature* **409:** 614–618.

Murphy, W.J., Eizirik, E., O'Brien, S.J., Madsen, O., Scally, M., Douady, C.J., Teeling, E., Ryder, O.A., Stanhope, M.J., de Jong, W.W., et al. 2001b. Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* **294:** 2348–2351.

Murphy, W.J., Pevzner, P., and O'Brien, S.J. 2004. Mammalian phylogenomics comes of age. *Trends Genet.* **20:** 631–639.

Nikaido, M., Nishihara, H., Hukumoto, Y., and Okada, N. 2003. Ancient SINEs from African endemic mammals. *Mol. Biol. Evol.* **20:** 522–527.

Nishihara, H., Satta, Y., Nikaido, M., Thewissen, J.G.M., Stanhope, M.J., and Okada, N. 2005. A retroposon analysis of Afrotherian phylogeny. *Mol. Biol. Evol.* **22:** 1823–1833.

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

Notredame, C., Higgins, D., and Heringa, J. 2000. T-Coffee: A novel method for multiple sequence alignments. *J. Mol. Biol.* **302:** 205–217.

Pecon-Slattery, J., Pearks-Wilkerson, A.J., Murphy, W.J., and O'Brien, S.J. 2004. Phylogenetic assessment of introns and SINEs within the Y chromosome using the cat family Felidae as a species tree. *Mol. Biol. Evol.* **21:** 2299–2309.

Poux, C., van Rheede, T., Madsen, O., and de Jong, W.W. 2002. Sequence gaps join mice and men: Phylogenetic evidence from deletions in two proteins. *Mol. Biol. Evol.* **19:** 2035–2037.

Roca, A., Bar-Gal, G.-K., Eizirik, E., Helgen, K.M., Maria, R., Springer, M.S., O'Brien, S.J., and Murphy, W.J. 2004. Mesozoic origin for West Indian insectivores. *Nature* **429:** 649–651.

Rokas, A. and Carroll, S.B. 2006. Bushes in the tree of life. *PLoS Biol.* **4:** e352.

Rokas, A. and Holland, P.W.H. 2000. Rare genomic changes as a tool for phylogenetics. *Trends Ecol. Evol.* **15:** 454–459.

Rosenbloom, K., Taylor, J., Schaeffer, S., Kent, J., Haussler, D., and Miller, W. Phylogenomic resources at the UCSC Genome Browser. In *Methods in molecular biology: Phylogenomics* (ed. W.J. Murphy). Humana Press, Totowa, NJ. (in press).

Schwartz, M., Kent, W.J., Smit, A., Zhang, Z., Baertsch, R., Hardison, R.C., Haussler, D., and Miller, W. 2003. Human-mouse alignments with BLASTZ. *Genome Res.* **13:** 103–107.

Shedlock, A.M. and Okada, N. 2000. SINE insertions: Powerful tools for molecular systematics. *Bioessays* **28:** 148–160.

Shedlock, A.M., Takahashi, K., and Okada, N. 2004. SINEs of speciation: Tracking lineages with retroposons. *Trends Ecol. Evol.* **19:** 545–553.

Shoshani, J. and McKenna, M.C. 1998. Higher taxonomic relationships among extant mammals based on morphology, with selected comparisons of results from molecular data. *Mol. Phylogenet. Evol.* **9:** 572–584.

Smith, A.G., Smith, D.G., and Funnell, B.M. 1994. *Atlas of Cenozoic and Mesozoic coastlines*. Cambridge University Press, Cambridge, UK.

Springer, M.S., Murphy, W.J., Eizirik, E., and O'Brien, S.J. 2003. Placental mammal diversification and the Cretaceous-Tertiary boundary. *Proc. Natl. Acad. Sci.* **100:** 1056–1061.

Springer, M.S., Stanhope, M.J., Madsen, O., and de Jong, W.W. 2004. Molecules consolidate the placental mammal tree. *Trends Ecol. Evol.* **19:** 430–438.

Thomas, J.W., Touchman, J.W., Blakesley, R.W., Bouffard, G.G., Beckstrom-Sternberg, S.M., Margulies, E.H., Blanchette, M., Siepel, A.C., Thomas, P.J., McDowell, J.C., et al. 2003. Comparative analyses of multi-species sequences from targeted genomic regions. *Nature* **424:** 788–793.

Thorne, J.L., Kishino, H., and Painter, L.S. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* **15:** 1647–1657.

van de Lagemaat, L.N., Gagnier, L., Medstrand, P., and Mager, D.L. 2005. Genomic deletions and precise removal of transposable elements mediated by short identical DNA segments in primates. *Genome Res.* **15:** 1243–1249.

van Rheede, T., Bastiaans, T., Boone, D.N., Hedges, S.B., de Jong, W.W., and Madsen, O. 2006. The platypus is in its place: Nuclear genes and indels confirm the sister group relation of monotremes and therians. *Mol. Biol. Evol.* **23:** 587–597.

Waddell, P.J. and Shelley, S. 2003. Evaluating placental inter-ordinal phylogenies with novel sequences including RAG1, gamma-fibrinogen, ND6, and mt-tRNA, plus MCMC-driven nucleotide, amino acid, and codon models. *Mol. Phylogenet. Evol.* **28:** 197–224.

Waddell, P., Okada, N., and Hasegawa, M. 1999. Toward resolving the inter-ordinal relationships of placental mammals. *Syst. Biol.* **48:** 1–5.

Waddell, P.J., Kishino, H., and Ota, R. 2001. A phylogenetic foundation for comparative mammalian genomics. *Genome Inform. Ser. Workshop Genome Inform.* **12:** 141–154.

Wheeler, W.A. 2003. Implied alignment: A synapomorphy-based multiple-sequence alignment method and its use in cladogram search. *Cladistics* **19:** 261–268.

Wilson, M.D., Riemer, C., Martindale, D., Schnupf, P., Boright, A., Cheung, T., Hardy, D., Schwartz, S., Scherer, S., Tsui, L.C., et al. 2001. Comparative analysis of the gene dense *ACHE/TFR2* region on human chromosome 7q22 with the orthologous region on mouse chromosome 5. *Nucleic Acids Res.* **29:** 1352–1365.

Woodburne, M.O., Rich, T.H., and Springer, M.S. 2003. The evolution of tribospheny and the antiquity of mammalian clades. *Mol. Phylogenet. Evol.* **28:** 360–385.

Yu, L. and Zhang, Y.P. 2005. Evolutionary implications of multiple SINE insertions in an intronic region from diverse mammals. *Mamm. Genome* **16:** 651–660.

Received September 1, 2006; accepted in revised form December 20, 2006.

Using genomic data to unravel the root of the placental mammal phylogeny

William J. Murphy, Thomas H. Pringle, Tess A. Crider, et al.

Genome Res. 2007 17: 413-421 originally published online February 23, 2007
Access the most recent version at doi:10.1101/gr.5918807