Cytogenet Genome Res 108:26–37 (2005)
DOI: 10.1159/000080799

Cytogenetic and Genome Research

Primate phylogeny: molecular evidence from retroposons

J. Schmitz,[a] C. Roos[b] and H. Zischler[c]

[a] Institute of Experimental Pathology (ZMBE), University of Muenster, Muenster;
[b] Primate Genetics, German Primate Centre, Goettingen;
[c] Institute of Anthropology, Johannes Gutenberg University, Mainz (Germany)

Abstract. In these postgenomic times where aspects of functional genetics and character evolution form a focal point of human-mouse comparative research, primate phylogenetic research gained a widespread interest in evolutionary biology. Nevertheless, it also remains a controversial subject. Despite the surge in available primate sequences and corresponding phylogenetic interpretations, primate origins as well as several branching events in primate divergence are far from settled. The analysis of SINEs – short interspersed elements – as molecular cladistic markers represents a particularly interesting complement to sequence data. The following summarizes and discusses potential applications of this new approach in molecular phylogeny and outlines main results obtained with SINEs in the context of primate evolutionary research. Another molecular cladistic marker linking the tarsier with the anthropoid primates is also presented. This eliminates any possibility of confounding phylogenetic interpretations through lineage sorting phenomena and makes use of a new point of view in settling the phylogenetic relationships of the primate infraorders.

Primates, our own eutherian order, constitute a remarkably diverse taxon. With at least 230 species recognized at present (Rowe, 1996), primates appear to be one of the most successful eutherian orders, surpassed in species number only by rodents, chiropterans, insectivores and carnivores. Due to ongoing and intensified field-work, the continuous generation of primate sequence data and the application of genetic data in primate taxonomy, evidence is mounting that the species diversity of extant primates is actually greatly underestimated. Indeed, several authorities in primate research propose an extension of the primate species list to over 350 (see e.g. Groves, 2001; Geissmann, 2002), leading to an anticipated increase in primate biodiversity in the short term. This however should not detract from the fact that many primate taxa are of considerable conservation concern or on the fringe of extinction due to habitat destruction and hunting.

Recent compilations of extended molecular data sets – both mitogenomic and nuclear types – suggest that extant eutherians can be partitioned into four major groups, the Laurasiatheria, Xenarthra, Afrotheria and the Euarchontoglires (Madsen et al., 2001; Murphy et al., 2001a, b). Together with the orders of rodents, lagomorphs, scandentians and dermopterans, primates are members of the latter.

Most experts consider rodents and lagomorphs to represent monophyletic sister groups. The phylogenetic affiliations among the remaining members of the Euarchontoglires, the primates, flying lemurs (Dermoptera) and tree shrews (Scandentia) however remain highly disputed. This is all the more surprising since complete genome sequences from humans and the mouse as the prime mammalian model are now at hand, marking the beginning of the postgenomic era and setting the stage for a central role of human-mouse comparisons in biomedical science. In this context, the effect of human-mouse sequence differences on transcriptomes and proteomes, the increase in transcriptome and proteome complexity on the lineage to

Supported by the DFG and by the European Commission (INPRIMAT, contract QLRI-2002-01325).

Received 24 October 2003; revision accepted 6 February 2004.

Request reprints from: Hans Zischler, Institute of Anthropology
Johannes Gutenberg University, Saarstrasse 22
D–55099 Mainz (Germany); telephone: +49 6131 39 24354
fax: +49 6131 39 25132; e-mail: zischler@mail.uni-mainz.de

0301–0171/05/1083–0026$22.00/0

Accessible online at:
www.karger.com/cgr

humans through alternative splicing, RNA editing, epigenetic and posttranslational modifications for example are just a few likely candidates for future biomedical research. Knowledge of the phenomena that shape the spatio-temporal pattern of transcriptomes and proteomes during the development of organisms is hoped to contribute to our understanding of what makes us human.

Any analysis of character evolution, whether focusing on morphological, molecular, physiological or behavioral characters, requires a solid phylogenetic framework in order to discriminate between homology and analogy. Since non-human primates are the closest relatives of *Homo sapiens*, it is obvious that the phylogenetic history of primates and their closest eutherian relatives provides a link to the mouse as a model organism. Their phylogenetic history thus forms part of our own evolutionary history.

A prime objective of retroposon research in primates and their closest relatives is to unravel their phylogenetic relationships. This is all the more important since a closer look at primate evolutionary research reveals several discrepancies between interpretations based on neontological and palaeontological morphological data. Discrepancies are also apparent on a molecular level between mitochondrial (mt) DNA- and nuclear (nuc) DNA-based data. Due to their entirely different approach, retropositional analyses allow alternative topology proposals to be tested, providing a fresh approach to gridlocked phylogenetic problems that seem to resist all available sequence data.

At the same time, phylogenetic research on the "pattern" clearly represents only one aspect of retroposon research in primates and primate-related taxa. Considering the sheer number of retroposons already transpositionally quiet or still showing evidence of mobility, it is possible that these sequences and their steady emergence have considerably influenced genome function, plasticity and architecture over evolutionary time and continue to do so at present (for review see Brosius, 1999).

Retroposons

As typical eutherian genomes primate genomes are peppered with a variety of repetitive sequences that form a major part of the respective genome but display no global function (Smit and Riggs, 1995). The bulk of these discernible sequences, constituting more than 40% of the human genome (Nekrutenko and Li, 2001), represents remnants of transpositional activity of mobile molecules that were major factors in shaping the genomes of extant taxa during their evolution. Mobile sequences can move from their parent genomic location to their target either by a cut and paste mechanism involving a DNA intermediate, or they transpose replicatively by a copy and paste mechanism via an RNA-intermediate (retroelements). The former are usually designated as class II elements. Class I or retroelements can be further partitioned into major groups, for instance elements that code for their own reverse transcriptase, thus allowing autonomous transposition. This group includes the non-LTR (long terminal repeat) elements such as LINEs (long interspersed elements) and LTR elements as well as the true retroviruses. Short interspersed elements (SINEs) form another type of elements in this group with non-autonomous transposition, exploiting instead the enzymatic retropositionary machinery of other elements such as LINEs. During their retroposition, the RNA intermediate is reverse transcribed by a process called target primed reverse transcription (TPRT). The complete reverse transcription and reintegration process is apparently mediated by factors provided by the LINEs which encode both proteins with reverse transcriptase and endonuclease activity (Luan et al., 1993; Feng et al., 1996; Jurka, 1997; Kazazian and Moran, 1998; Kajikawa and Okada, 2002).

Repetitive elements and all types of transposition clearly contributed to the configuration and plasticity of the eutherian genomes. Investigations of retropositions as phylogenetic markers must therefore bear in mind that all above phenomena might have had functional implications during the process of genome construction, for instance by generating novel protein-coding domains or modulating genes and their spatio-temporal expression (reviewed in Brosius, 1999).

In this review, we will limit ourselves to the "pattern" problems of evolutionary research and the question of how SINEs can be used to test phylogenetic hypotheses. The sole focus on SINEs as transposable markers is simply due to the fact that these are most commonly used in phylogenetic studies. Theoretically, other elements that transpose replicatively and are excluded from horizontal transfer could equally serve as phylogenetic marker systems, for instance non-LTR elements (shown in Malik et al., 1999). Another reason for favoring SINEs in phylogenetic examinations is that short SINE sequences can easily be checked for presence/absence at orthologous loci. In addition, the high copy number of SINEs generally available in a typical eutherian represents an almost inexhaustible source of information on the evolutionary history of a certain genome.

To date, LINE-based phylogenetic analyses mostly draw on presence/absence analyses of complete LINE-families or the use of LINE-based sequence information in certain taxa (Verneau et al., 1998). Comparative analyses of orthologous endogenous retroviruses, integrated retroviral sequences or their remnants were mostly used to show the evolutionary history of the viral sequence. Rarely were they employed as a phylogenetic marker system for their hosts (Sverdlov, 2000).

SINEs appear to be the most abundant class I sequence to be traced in eutherian genomes. They typically range between 75 and 500 bases in size and may be amplified to a copy number well beyond $>10^4$ total copies per genome. The majority of the eukaryotic SINE families described so far can be traced back to a tRNA ancestor, with a minority apparently derived from 7SL RNA. A common feature of both types of SINEs are their internal promoter sequences which are specific for RNA polymerase III (Okada, 1991a, b).

Primates and some primate-related taxa harbor both types of SINEs in their genomes (see below). However, the most successful primate-specific SINE in terms of broad taxonomic distribution and copy number per genome is the so-called Alu SINE with an approximate full length of 300 bp. Over one million Alu sequences have been identified in the human genome, comprising about 10% of its overall sequence (Li et al., 2001).

Alu copy number estimates for other primate species are rare, however great apes and strepsirrhine representatives are estimated to harbor several hundred thousands of Alu copies in their genomes as well. The typical dimeric Alu repeats are primate specific. After dimerization from so-called free monomers, they expanded in number in a wave-like fashion during primate divergence. Alu sequences are frequently found in introns, untranslated regions of genes and intergenic genomic regions (Deininger and Batzer, 1993; Makalowski et al., 1994).

Available data on the taxonomic distribution of all eukaryotic SINE families described to date suggests that specific SINE families originated and amplified in particular taxonomic groups. These families can be further divided into subfamilies on the basis of diagnostic substitutions. Like all other mammals, primates for instance harbor so-called MIR sequences, which are assumed to have spread through the genome of a common ancestor that predated the current mammalian divergence.

After reintegration into the host genome, SINE sequences are assumed to evolve neutrally. Paralogous MIR sequences in the human genome are therefore highly divergent, suggesting coalescence on an element that was active early in mammalian evolution. On the other hand, primates also harbor the Alu sequences mentioned above, whose greater similarity suggests a more recent origin of the primate Alu element. Indeed, the origin of the typical primate Alu sequence is correlated with the emergence of the primate order. Although Alu sequences closely resemble one another, they can be further subdivided into different subfamilies based on diagnostic mutations. This might be explained by the possibility that only a small subset of Alu elements is retropositionally active for a certain period, thus serving as an Alu source or "master" gene (Deininger et al., 1992). Two different models of SINE evolution were proposed, both referring to the retropositional activity of offspring SINEs. On the one side the master gene hypothesis predicts that a few active "master" SINEs are capable of retropositional amplification (Deininger et al., 1992), whereas the "multiple source gene model" forecasts that the progeny SINE copies can potentially propagate at the same extent as the parent SINEs they are derived from (Schmid and Maraia, 1992). Different models apparently apply to different SINEs, with primate Alu-sequences being discussed to follow the source gene model (reviewed in Shedlock and Okada, 2000). We speculate that certain SINEs can also follow a mode of dispersal that is intermediate between these two extreme assumptions. Alu-source genes accumulate diagnostic new mutations over time and establish a set of Alu subfamilies that retroposed in chronological order, starting with the oldest (Alu-J), leading on through the intermediate (Alu-S) and ending with the youngest (Alu-Y) (Batzer et al., 1996). Although some older Alu families might still be active at a very low level (Johanning et al., 2003), the majority of Alu elements sticks to this chronological order of transpositional activity. Some of the young Alu elements described for humans have been integrated at their genomic location so recently that they have not become fixed and are absent at the respective orthologous sites in great apes (Batzer and Deininger, 1991; Perna et al., 1992; Batzer et al., 1994).

SINEs as cladistic markers

Much research has been directed at the biology, genetics and genome shaping as well as functional implications of retroposable elements. As temporal landmarks of evolution, SINEs have also gained widespread application in evolutionary biology. The reason for this is that SINEs are assumed to represent powerful noise-free Hennigian synapomorphies as formulated by Shedlock and Okada (2000).

Considering the relatively unspecified targets and the size of a typical primate nuclear genome, the chance of SINE sequence integrations independently involving the same targets is negligible even over evolutionary time scales. Moreover, SINE integrations are assumed to be irreversible events since no biological mechanisms have yet been described for the precise re-excision of class I transposons. A clear differentiation between ancestral and derived character state at the respective locus thus becomes possible. Both features – the virtual lack of homoplasies combined with a clear character polarity – render SINE integration markers ideal tools for determining the common ancestry of two taxa by a shared derived transpositional event (Hamdi et al., 1999; Shedlock and Okada, 2000). Since both features of SINE integrations are relevant for the application of SINE markers in evolutionary research, a few more details concerning these properties will be exemplified for the most abundant primate SINE, the Alu sequence, in greater detail. Although Alu repeats may preferentially integrate into locally AT-rich regions in chromosome R bands (Korenberg and Rykowski, 1988; Matera et al., 1990) or might have a tendency to prefer integration into DNA regions that can adopt alternative structures, for instance kinks, no unambiguous target sequence or hotspot has yet been reported for Alu integrations based on comparing the integration flanking-repeats representing the target sites of integration (Jurka et al., 1998) (see below).

Although an increase in available comparative data might lead to the definition of some rare hotspot sites, as was the case in completely unrelated SINE-LINE combinations in rodents (Cantrell et al., 2001), it is reasonably safe to assume that SINE markers will still display far lower convergence frequencies and parallelisms than sequence data and complex morphological characters. Alu insertions, present at the orthologous positions in different genomes, can thus be reliably considered identical by descent.

Another major factor pointing to a true orthology of Alu elements are their flanking direct repeats. The presence of short direct repeats flanking the SINE suggests integration in the target genome via staggered end breaks. These might result from endonucleolytic enzyme activity that can be attributed to a LINE-encoded endonuclease mediating the reintegration into the nuclear genome (Feng et al., 1996; Jurka, 1997). The unduplicated repeat sequence can thus be considered a target of the integration, which means that even minute differences in the integration sites can theoretically be pinpointed (Salem et al., 2003a, b). As a last point, the sequence information of the repeat unit itself might be sufficient to clearly assign the Alu element to a certain subfamily. Identity of the Alu subfamily is a further sign of true orthology, although rare gene conversion

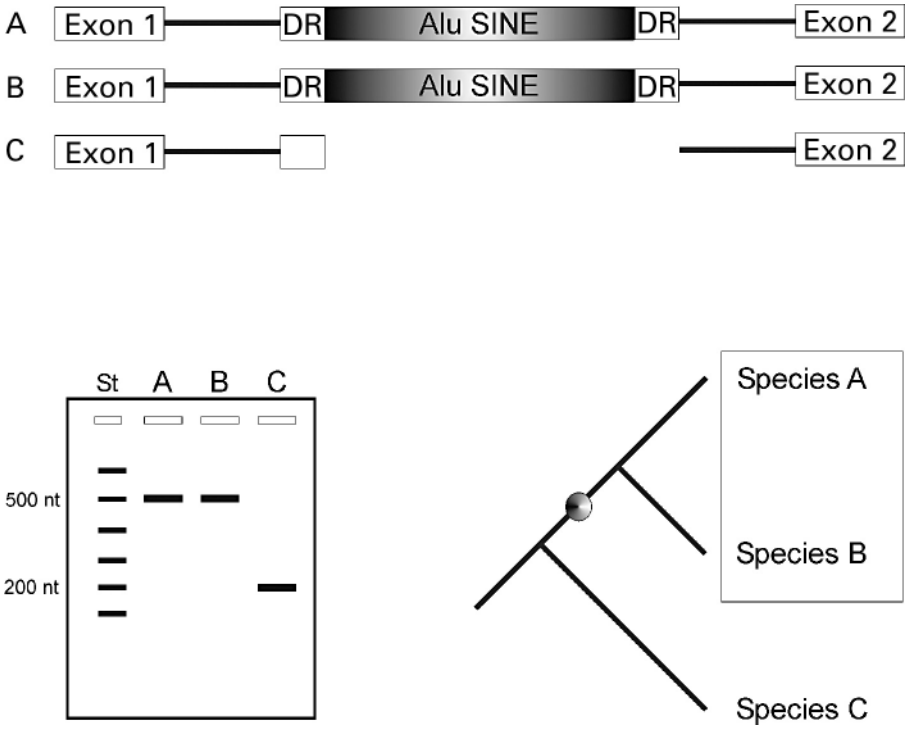

Fig. 1. Schematic representation of the experimental strategy used to analyze the presence/absence pattern of intronic Alu sequences at orthologous loci. Primers are placed in conserved exonic regions to check for intronic Alu elements that are characterized by a fragment size increase of about 300 bp (in A and B). Note, that the absence of an orthologous Alu-element is characterized by a fragment size shift of the respective PCR product of about 300 bp (C). Sequencing of the products should reveal Alu-flanking direct repeats (DR) and the respective sequence in unduplicated form represents the unoccupied insertion target sequence (open box for species C). The most parsimonious scenario of the Alu integration is displayed as well. St denotes the size standard.

events might act to confound interpretations based on subfamily diagnoses (Salem et al., 2003a, b).

A further huge advantage of Alu integrations compared to sequence data for instance is the complete absence of homoplasies resulting from character reversals. Since it is virtually impossible to precisely re-excise an Alu element, discriminating between the presence of an Alu insertion (reflecting the derived character state) and an unoccupied target site (reflecting the ancestral state) is just a matter of technique. The latter state is hallmarked by the presence of the target site which is the repeat sequence next to the SINE integration in unduplicated form. In technical terms, PCR primers need to be placed well away from the Alu element and some single-copy flank. A shorter fragment thus signals the absence of an Alu element, while a fragment ca. 300 bp longer signals presence (Fig. 1).

Naturally, SINE-based molecular cladistic analyses also have some limitations. One problem is that most SINEs are located in non-coding regions, regions which usually display considerable genetic distances in the context of analyzing deeper splits in primate divergence. As a consequence, PCR primers might not possess targets sufficiently conserved to allow efficient amplification across highly divergent taxa. To counter this, Schmitz et al. (2001) used intronic SINEs as markers that allow the placement of primers in conserved exonic regions, thus amplifying the encompassed intron together with the SINE. Successful amplification at least can be obtained for probably all members of an order, with cross-order amplification as an extended possibility.

One major disadvantage compared to conventional sequence analyses is that presence/absence analyses of Alu elements are purely cladistic, meaning they do not inform on the timing of splitting events. However, this entirely different approach renders molecular cladistic analyses an ideal complement to classic sequence analyses for example. The purely cladistic nature of Alu transposition does not mean that temporal correlations are altogether lacking. As outlined above, most Alu elements can be classed into subfamilies – categorized by diagnostic mutations – each of which has its own time frame of retropositional activity. After being transposed, SINEs stay in place and decay according to a neutral mode of sequence evolution. This means that Alu elements active millions of years ago should be able to resolve phylogenetic splittings that took place during that same period of time. Alus transposed very recently can thus be used to reconstruct human infraspecific phylogenies, whilst older Alu-subfamilies are suitable for investigating the deep splits in primate phylogeny, e.g. the affiliation of tarsiers to other primate groups.

Another problem affecting every polymorphic marker system is the uneven distribution of ancestral polymorphisms into progeny lineages, a phenomenon termed incomplete lineage sorting. This affects very recent splits as well as deeper, consecutive splitting events that took place within a short period of time. Since the effective population sizes accompanying a splitting event are mostly unknown, it is difficult to determine whether a SINE marker was fixed in the ancestral progenitor population at the base of the consecutive split. It is therefore advisable to collect split-specific SINE information at multiple independent loci in order to test for any inconsistencies in marker interpretations that might have been caused by lineage sorting (Takahashi et al., 2001). Complete human sequence information and huge data sets from other primates now allow

the characterization of sufficient numbers of SINE markers to avoid errors in this respect.

Murata and coworkers (1993) first demonstrated the potential of retroposons as phylogenetic markers to uncover salmonid phylogenetic relationships. Primate SINE analysis in the context of primate evolutionary biology was first suggested by Ryan and Dugaiczyk (1989). Since then, the seminal work of the Okada group mainly working on Cetacea and salmonids has contributed much in other taxonomic groups to this field. A short description of the current state of research within and around primates is given below. For the sake of clarity the remaining phylogenetic questions in primate and primate-related taxa will be summarized in more or less "chronological" order, using one possible splitting pattern of primates and their closest relatives as a guide.

Primate origins and affiliation to other eutherians

Primates are characterized by a long list of common morphological features (Napier and Napier, 1967) such as primitive pentadactyly, enhanced mobility of the digits, nail development and the evolution of full stereoscopic vision. The definition of what actually constitutes a primate and separates it from other eutherians however is a problematic and complex issue. No unique morphological, physiological or behavioral character appears to exist that can be regarded as an undisputed synapomorphy, thus defining a clade to comprise all primate taxa. Many of the "typical primate" morphological characteristics represent retentions of ancestral features, and many "typical primate" features are behavioral or depend on soft tissue anatomy. Identifying a fossil and assigning it to the primates is thus a delicate task. Earliest primates are only marginally different from the representatives of other eutherian orders. Based on fossil evidence, the origin of extant primates is currently traced back to two species of the archaic primate genus *Purgatorius* which lived in North America and Eurasia at least 65 MYA during the late Cretaceous and early Paleocene (Shoshani et al., 1996).

A major concern with fossil evidence and primate origins and early evolution is the low sampling level. Tavaré and coworkers (2002) suggest that the fossil record only shows 7% of all primate species that ever existed. Taking this into account and applying a new statistical method that uses diversification patterns to estimate species preservation, they estimate that the most recent common ancestor (MRCA) of all primates existed about 81.5 million years ago (MYA). This is considerably older than previous estimates based on *Purgatorius* fossils for instance.

Evidence is also mounting that the archaic primates or Plesiadapiformes are much more diverse than previously assumed. Various phylogenetic constellations have been proposed and are still being discussed for the different plesiadapiform families, including the Micromomyidae, Saxonellidae, Carpolestidae, Plesiadapidae, Picrodontidae and Paromomyidae and extant members of the Archonta superorder, i.e. the euarchontan orders Primates, Dermoptera, Scandentia and Chiroptera (reviewed in Martin, 1993). Molecular data obtained from extant taxa provides valuable assistance in forming theories on the phylogenetic affiliations between primate and non-primate crown groups and can help to direct the development of evolutionary hypotheses on the archaic primates.

Currently, the two eutherian orders of Dermoptera (flying lemurs) and Scandentia (tree shrews) are being discussed as potential extant sister groups to the primates. Molecular evidence – both nuclear and mitogenomic – supporting an exclusion of bats from a close alliance to primates is now strong (Pumo et al., 1998; Teeling et al., 2000). The supposed close phylogenetic relationship between Dermoptera and Chiroptera is probably based on convergences that emerged with the evolution of gliding or flight.

Indeed, Beard (1990, 1993) observed similarities in postcranial features and phalangeal morphology of paromomyids and extant taxa, suggesting that some genera of the Plesiadapiformes might be more directly linked to Dermoptera. He also suggests that the former possessed a gliding membrane (patagium) similar to extant flying lemurs (for an alternative discussion see Kay et al., 1990).

Sequence data from flying lemurs are at hand, but the phylogenetic results obtained from these are contradictory. This remains the case even when different taxonomic samplings are taken into account.

Most unexpected to primatologists was the close affiliation of dermopterans to anthropoid primates, which was proposed from a composed mtDNA and nuclear data set (Murphy et al., 2001a) and through complete mtDNA information alone (Arnason et al., 2002). The term "Dermosimii" was introduced to denote the close phylogenetic relationship between anthropoid primates and the dermopterans. This effectively rendered the primates a paraphyletic group.

This complex problem is a prime example for how SINE evidence can help to test conflicting phylogenetic hypotheses. Indeed, the above constellation is not corroborated by SINE data. On a multi-locus level, Schmitz and co-workers (2002b) were able to show that dermopterans have no or at least an enormously reduced copy number of Alu sequences compared with primates.

Naturally, the mono-locus presence/absence analysis of SINEs at orthologous loci has all the benefits of a clear underlying formal genetic model. However, the multi-locus approach with its simultaneous analysis of virtually all paralogous Alu sequences of a taxon's genome is a quick and simple experimental procedure that does not require any additional sequence information other than the SINE sequence itself. Hybridization with a specific SINE probe allows a relatively precise assessment of the presence of the respective SINE in a taxonomic group, permitting even a rough estimate of the respective SINE copy number.

In the topology proposed by Arnason et al. (2002) this would stipulate a simultaneous elimination of hundreds of thousands of Alu copies from the genome of dermopterans once they split off from the lineage leading to the ancestor of anthropoid primates. Alu hybridization data therefore agree with the long-held perception that the spreading of dimeric Alu elements is restricted to primates. On the other hand, single loci bearing Alu elements observable in tarsiers could not be traced in der-

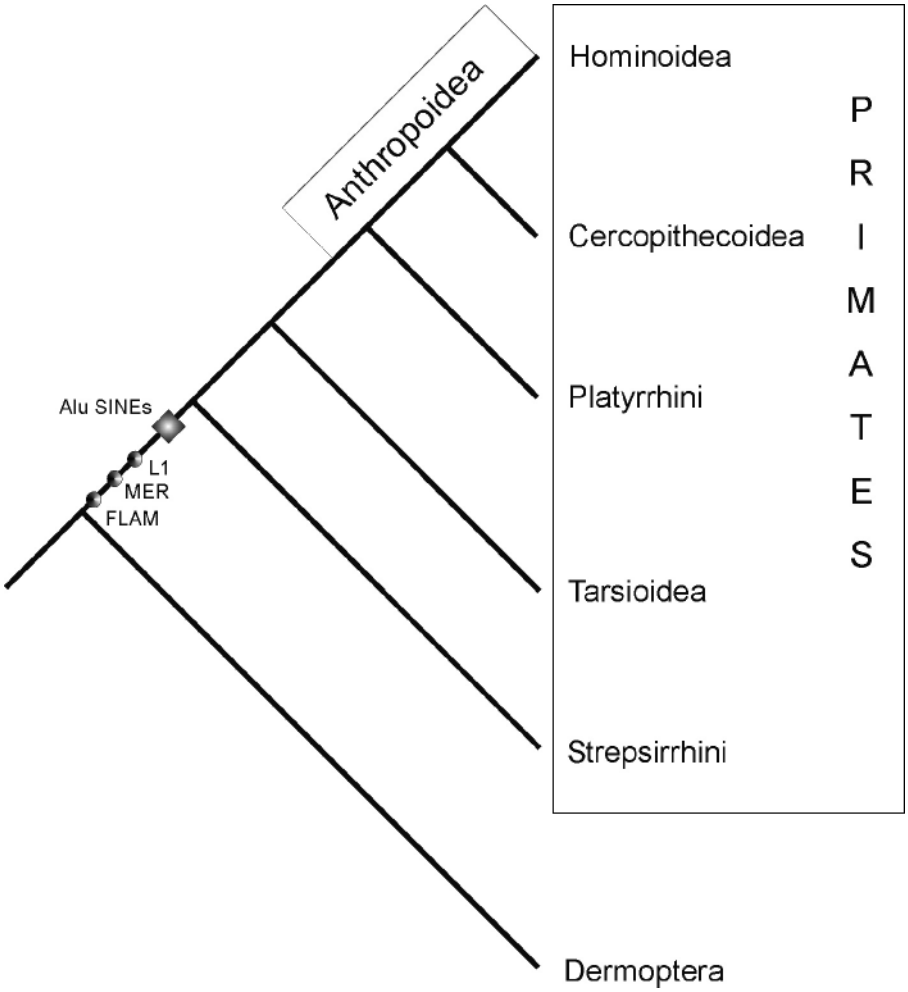

Fig. 2. The "Dermosimii" problematics revisited. Both multilocus (Alu-SINEs) as well as single locus evidence point towards a clear monophyly of primates thus excluding the postulated close phylogenetic relationship of the Dermoptera to anthropoid primates based on mitogenomic information. Strepsirrhini and Tarsioidea designate the representatives of the two deepest splits in primate divergence.

mopterans, whose orthologous loci instead display the ancestral character state of an unoccupied target (Schmitz et al., 2002a). Moreover a set of three monolocus markers that are shown by representatives of all primate infraorders to the exclusion of other, non-primate eutherians including dermopterans (Schmitz et al., 2002b; Schmitz and Zischler, 2003) could be defined thus corroborating primate monophyly from a retropositional site (summarized in Fig. 2).

Accordingly, dermopterans still need to be considered a possible sister group of primates, although they do not share primate variability. This raises the question why mitochondrial DNA data for instance yield such an interpretation. Rather than reflecting true phylogenetic relationships, the unexpected positioning of the dermopterans is considered the result of a similar mtDNA base composition. This effect is assumed to be so strong that even combined mtDNA and nuclear data sets might be affected (Schmitz et al., 2002a, b).

Apart from the dermopterans, the members of the eutherian order Scandentia form another candidate sister group of primates. Living representatives of this order are generally subdivided into two subfamilies, the Tupaiinae and Ptilocercinae. Tupaiinae, or tree shrews, are actually a misnomer since they are not unvaryingly arboreal. They represent a small radiation of eutherians with 18 extant species currently recognized and a geographical range restricted to southern and south-east Asia (Martin, 1990). Varying phylogenetic affiliations attributed to this taxon indicate the uncertainty that continues to be associated with scandentian evolution. Taxa currently assigned to the Scandentia have previously been included in the insectivore subgroup Menotyphla or were recognized as the deepest primate split before constituting their own mammalian order of Scandentia (Martin, 1990).

Tupaias display a complex mixture of plesiomorphic and apomorphic morphological characters (Starck, 1978; Martin, 1990), hampering the definition of a clear phylogenetic position. The scarcity of fossils pertaining to the early primate emergence and divergence, the problem of "early primate definition" and the scanty fossil record accompanying tupaian evolution render the development of consistent hypotheses on euarchontan evolutionary history all the more difficult.

SINE information is expected to be a main factor in solving this question and verifying the different hypotheses on euarchontan affiliations.

With the exception of xenarthran species, different SINE families and subfamilies have been described in many lineages of the four major eutherian clades proposed e.g. in Murphy et

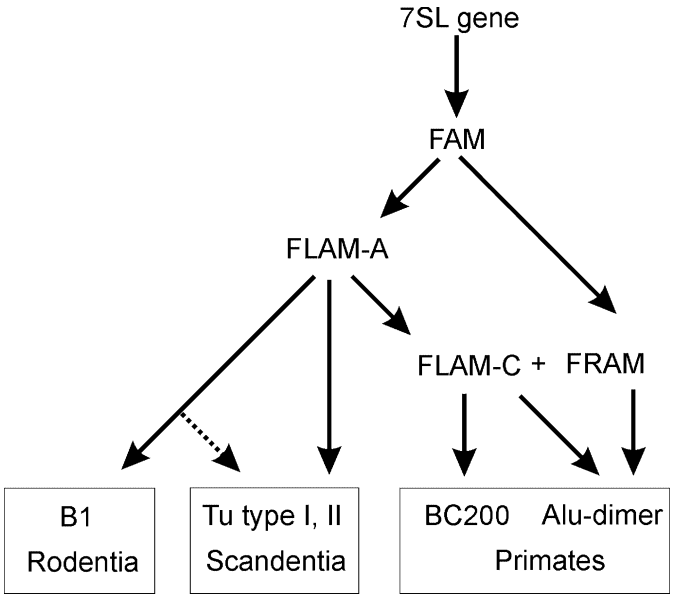

Fig. 3. Summary of the presently discussed evolutionary history of 7SL-derived SINEs in the Euarchontoglires group of eutherians. The dotted arrow indicates a possible close phylogenetic affiliation of rodents and the Scandentia, an issue that requires more data for a concluding appraisal.

al. (2001a, b) (see Shimamura et al., 1997, 1999; Nikaido et al., 2003). Current research therefore focuses on defining SINEs in those Euarchontoglires that constitute a phylogenetic link between the different orders. Interestingly, several Euarchontoglires harbor SINEs that are apparently derived from 7SL RNA, including the rodent B1, the primate Alu, and the scandentian Tu type I and II SINE families (Nishihara et al., 2002) (see also Fig. 3).

No 7SL RNA derived SINE family has yet been detected in rabbits or flying lemurs. The lagomorph genome harbors the C repeat SINE family (Cheng et al., 1984) which appears to be derived from tRNA genes (Sakamoto and Okada, 1985).

Recently, Schmitz and Zischler (2003) characterized tRNA-derived elements in the flying lemur genome, results which were corroborated by Piskurek et al. (2003). These tRNA-derived elements display different degrees of tandem reiteration and deceptively constitute the major class of SINEs in these genomes. Similar elements could not be traced in primate genomes for example. This is corroborated by the fact that family members display only little divergence among each other, which is indicative for their coalescence into a recent ancestor. Of the 7SL-derived SINEs, all recognized Alu SINE subfamilies are exclusively distributed throughout the primate genomes (Britten et al., 1988; Schmid, 1996). Primate Alu elements represent a dimer that originated by a fusion of the left and right monomer (FLAM and FRAM). An even more ancient monomer unit that forms part of the Alu family was named FAM (Quentin, 1992a, b, 1994). Because of the similarity of the FLAM sequence to a precursor sequence of the rodent B1 family, it is reasonable to assume that the precursors of primate Alus, rodent B1 and scandentian Tu elements, the FLAMs and FAMs, might have been created in a common Euarchontoglires ancestor, implying that these monomers were transpositionally active during these times. However, direct evidence for the existence of these monomer sequences in the genomes of tree shrews or flying lemurs is still lacking, which might also be explained by sequence divergence that took place in the long and independent evolutionary history of the euarchontan orders. A substantially different initial copy number and/or activity state of these elements after speciation is another explanation. Interestingly, both tupaias and some strepsirrhine taxa – the loriforms (see below) – harbor a compound 7SL and tRNA-derived SINE in their genomes (Tu type I and II and the galago type II element). These not only resemble each other in terms of compound make-up, but also in terms of the secondary structures of the respective tRNA-derived regions (Nishihara et al., 2002). These sequences were speculated to have emerged from monomer-building blocks present in a common Euarchontoglires ancestor. However, the challenge remains of setting appropriate criteria of parsimony for the several scenarios charting the origin and history of these elements. The same goes for other mutually exclusive evolutionary scenarios of 7SL-derived SINEs of Euarchontoglires that mainly hinge on the phylogenetic position of tupaias (Fig. 3).

Mono-locus data for the presence/absence of these elements in different orders are unlikely to be obtained, although the sequence divergence among the members of this group is considerable. This suggests a deep coalescence of the elements into an active progenitor sequence. The limited copy number (ca. 10^2) will probably not be sufficient for establishing informative mono-locus data able to elucidate the relationships between the different euarchontan orders. This is due to the fact that PCR strategies are required based on conserved primer targets, e.g. in exonic regions, in order to enable cross-order amplifications of integration targets which usually evolve rapidly without any selective pressure.

A picture thus emerges that although we learn a lot about their evolution, 7SL-derived SINEs can probably not be used for establishing sister group relationships among the different Euarchontoglires on the mono-locus level.

Alternative scenarios however could be verified by exploiting other SINEs abundantly represented in our genomes. Sequence comparisons among family members reveal that different SINE families coalesce into common ancestors which were active during different periods of eutherian or primate evolution. In this respect MIR sequences in the human genome for instance show a far more pronounced mutual substitution level, suggesting that the vast majority of active MIR sequences existed prior to primate divergence. It appears that MIRs are slowly numerically superseded by the Alu SINEs during primate divergence with the latter even establishing a different mode of retrotransposition (Kajikawa and Okada, 2002). Whilst these are highly informative for our recent history, the former should yield information on eutherian evolution preceding the most recent common ancestor of primates. Considering the copy number of these sequences and the availability of complete genomic sequences of mouse and humans, retropositional evidence for euarchontan affiliations is likely to become available in the near future.

SINEs and primate phylogeny

Less disagreement exists about the phylogenetic affiliations between and within the major primate groups. Probably, the most challenging problem involving the different primate infraorders is the positioning of the Tarsiiformes, primates with a current geographical distribution in Southeast Asia.

In line with the order of branching events though, the review will first consider the deepest primate split which gave rise to strepsirrhine primates.

This group is commonly divided into two or three infraorders, the Lemuriformes and the Chiromyiformes, geographically exclusively restricted to Madagascar, and the Lorisiformes found on the two continents Africa and Asia. Strepsirrhines form a remarkably diverse group of primates constituting ca. 20% of all living primate species. Current species distribution and accepted palaeocontinental reconstructions suggest an African origin of strepsirrhines with two subsequent migrations to Madagascar and Asia. The colonization of Madagascar was accompanied by a remarkable adaptive radiation resulting in the broad diversification of extant Lemuriformes. At least 33 extant lemuriform species are currently known to inhabit Madagascar, classified into the Lemuridae, Lepilemuridae, Indriidae, Cheirogaleidae and Daubentoniidae families. This is all the more surprising considering the small area lemurs inhabit on the Malagasy island. A long-standing, ongoing question is how often Madagascar was colonized by lemuriform ancestors. Apart from the subfossil lemurs, strepsirrhine ancestors are poorly documented in the fossil record. Also, both the timing and the physical crossing of the Mozambique channel from Africa to Madagascar remain much disputed. A significant oceanic barrier separated Madagascar from Africa and the Indian subcontinent at least 165 and 88 MYA respectively. Even those estimates that interpret the fossil record in favor of an earlier primate MRCA and an earlier diversification of extant clades (Tavaré et al., 2002) would still need to deal with this problem. Hard to imagine at any rate, the crossing of such a huge water barrier is difficult to conceive as a repetitive event occurring several times in the history of Malagasy primates.

Phylogenetic reconstructions based on morphological evidence however do not yield unequivocal proof of monophyly of Malagasy lemurs and therefore fail to specify a single colonization event. Some morphological features, related to the intracranial blood supply for instance are shared by both the Malagasy cheirogaleids and the African galagids. This has repeatedly led to suggestions for a closer phylogenetic affiliation between the Cheirogaleidae and the galagids (reviewed in Martin, 1990), a tree topology that would support at least two independent colonizations of Madagascar. Molecular data firmly place the cheirogaleids with the lemuriforms, a conclusion further corroborated by SINE evidence.

In terms of SINE evidence, the cheirogaleid-galagid problem mirrors the issue of the Dermoptera-Primate affiliation. The lorisiform genome, including that of the galagids, harbors a composite tRNA-7SL-derived SINE that cannot be traced in any other strepsirrhine outside the lorisiform group (Roos et al., 2004). This multi-locus SINE evidence rules out an affiliation of Malagasy cheirogaleids and African galagids: Hundreds of thousands of copies of the respective SINE would need to be eliminated from the genome of the cheirogaleid ancestors, clearly rendering this phylogenetic association virtually impossible.

Another problem regarding the monophyly of Malagasy lemurs is far more complicated and concerns the position of the aye-aye *(Daubentonia madagascariensis)*. On a morphological and molecular level, conclusive evidence for the aye-aye's position and its relation to other Malagasy lemurs is still lacking. Alternative topologies have been proposed showing the aye-aye at the base of all strepsirrhines, joined with the Indriidae or not fully resolved with respect to other strepsirrhines (e.g. Arnason et al., 1998).

A conclusive SINE analysis could assist in settling the ongoing debate about the phylogenetic affiliations of the classical Malagasy lemur families, the Cheirogaleidae, Indriidae, Lemuridae and Lepilemuridae. For this however, a presumed radiation-like branching pattern must be taken into account which took place shortly after the lemur ancestors colonized Madagascar, most likely after *Daubentonia* split off. In such a scenario, lineage sorting, the unequal distribution of ancestral polymorphisms into progeny lineages, is a likely obstacle for the generation of consistent interpretations based on presence/absence data of SINEs at orthologous loci. To solve this problem of the retropositional marker site, several SINE markers informative for a certain branch will be required.

On the intraordinal taxonomic level, the most striking cladistic problem in evolutionary primatology concerns the phylogenetic affiliations of *Tarsius* to other extant primates. Differing taxonomic categorizations reflect the uncertainty concerning the tarsier position in the primate phylogenetic tree. One model separates *Tarsius* from the simians and assigns it to the prosimian group together with lemurs and lorises. A second taxonomic grouping places *Tarsius* with the New World and Old World monkeys in a group called Haplorrhini, which is separate from the strepsirrhines, the lorisiform, lemuriform and chiromyiform primates. In classical terms, the problem is mostly due to a conflict between neontological and palaeontological morphological data. As far as the molecular site is concerned, strong discrepancies exist between interpretations based on mitochondrial or nuclear data sets. At first, deviations from a purely neutral mode of mtDNA-evolution were invoked to explain this apparent misplacing (Andrews et al., 1998). However, the unexpected positioning of *Tarsius* in an mtDNA-derived phylogenetic tree could be explained by the same phenomenon of base composition plasticity of mitochondrial DNA that was outlined above for the Dermoptera-Anthropoidea (Schmitz et al., 2002a, b). Most nuclear DNA data sets firmly place *Tarsius* as a sister to the anthropoid primates ("omomyid theory"), thus supporting the haplorrhine grouping. Analyses of fossil records favor alternative evolutionary tree topologies which either place *Tarsius* as a sister group to the Strepsirrhini ("adapid theory"), show *Tarsius* to branch off before the Anthropoidea-Strepsirrhini split or give rise to a polytomy involving all three taxa. The *Tarsius* position is a very delicate problem since *Tarsius* is the only surviving genus of a formerly diverse group of Eocene Tarsiiformes. A major obstacle in solving it is the long independent history of *Tarsius* and

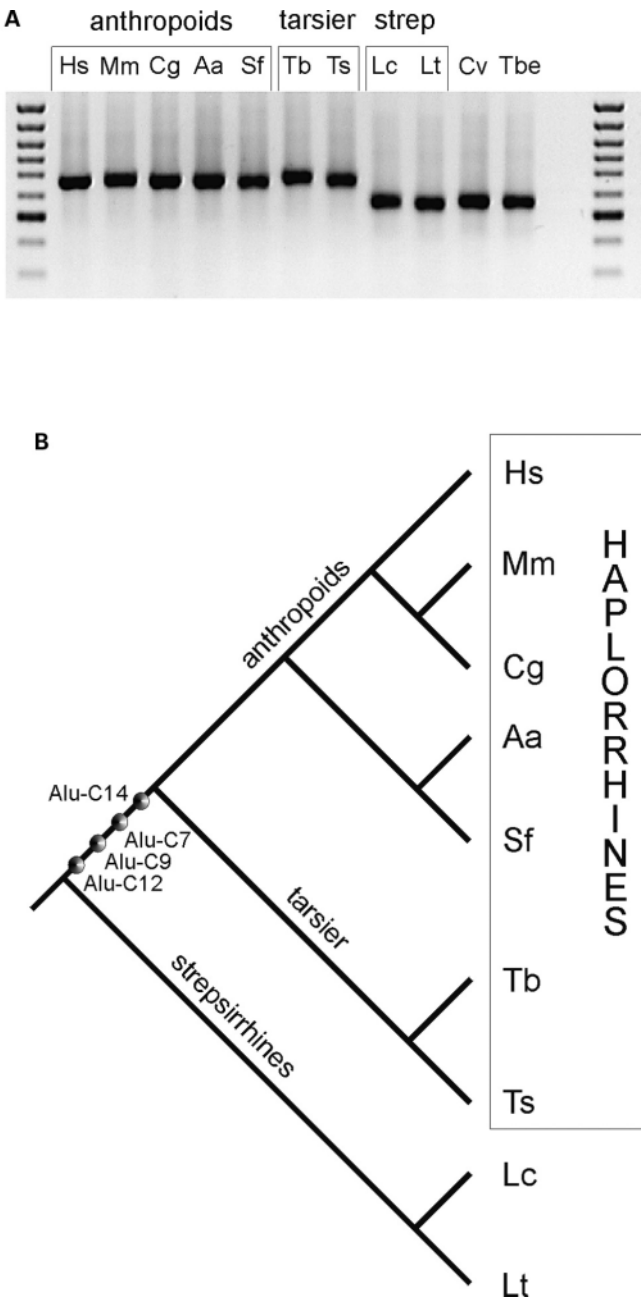

Fig. 4. (**A**) Alu sequences located in intronic regions were determined in database searches. Markers were chosen applying the criteria of available mouse sequences and manageable PCR fragment size. The SINE marker presented here is located in the serine palmitoyl transferase, subunit II gene on human chromosome 14 (AF111168) and defined as a FLAM A sequence (free left Alu monomer). Exonic primers (5′ CTG GTG GAA GAT GTG GAC AC and 5′ TRT CTA CCT TAC TCC TGT ATG C) were applied in standard hot start PCRs with 30 cycles, each consisting of 30 s at 94 °C, 30 s at 57 °C, and 60 s at 72 °C. The PCR fragments obtained from Tbe: *Tupaia belangeri*, Cv: *Cynocephalus variegatus*, Lt: *Loris tardigradus*, Lc: *Lemur catta*, Ts: *Tarsius syrichta*, Tb: *Tarsius bancanus*, Sf: *Saguinus fuscicollis*, Aa: *Aotus azarae*, Cg: *Colobus guereza*, Mm: *Macaca mulatta* and Hs: *Homo sapiens* (from right to left) were electrophoresed, ethidium bromide stained and are displayed with markers on both adjacent lanes. The taxa belonging to the anthropoid primates, tarsiers and the strepsirrhines (strep) are marked. Sequences were obtained from all taxa shown here, with the following respective accession numbers: AY388632-41. (**B**) Schematic drawing of the phylogenetic affiliations of the main primate groups. Alu-C7, Alu-C9 and Alu-C12 designate the most parsimonious integration time span for 3 Alu elements previously described (Schmitz et al., 2001). Alu-C14 designates the marker presented here. Altogether 4 different SINE markers clearly unite tarsiers and anthropoid primates in a monophylum of haplorrhines to the exclusion of the strepsirrhines and non-primate outgroups (not shown).

its high likelihood of acquiring autapomorphies, compounded by the fact that this one genus cannot properly reflect the complete tarsiiform diversity that once existed.

SINE evidence for the *Tarsius* position has so far been obtained using two very different approaches. The first was based on sequence comparisons of primate Alus, including Alu sequences from *Tarsius.* This approach resulted in a positioning of *Tarsius* as a sister to the anthropoid primates (Zietkiewicz et al., 1999). In another SINE analysis, Schmitz et al. (2001) analyzed the presence/absence pattern of 118 human loci containing intronic Alu sequences. Figure 1 displays the basic methodology of this approach.

Three Alu SINEs were found to be present at orthologous sites in *Tarsius* and all anthropoids. The respective unoccupied integration target sites were found in all strepsirrhine representatives and the non-primate outgroups. A degree of mutual dependence was considered for these analyses (Yoder, 2003), alleviating the phylogenetic information on this SINE-based approaches. However, these two analyses may not be comparable since the first relies on a multi-locus approach and analyzes paralogous sequences at the same time. Presence/absence information however is a single-locus approach that makes use of all the advantages provided by a well-defined formal genetic model. Further retropositional evidence for a close affiliation of anthropoids and *Tarsius* was also provided by Kuryshev et al. (2001). All data on retrotransposons support the haplorrhine scenario, with only lineage sorting phenomena constituting a possible interference. We therefore went on to characterize more human loci for the presence of intronic SINEs, focusing on those that could be checked for their presence/absence pattern throughout primate divergence. In this way we were able to define an intronic SINE in the serine palmitoyl transferase, subunit II gene on human chromosome 14, represented by a FLAM A sequence. Figure 4A displays the banding pattern obtained after electrophoresing the respective PCR products applying exonic primers that encompass the SINE-containing intronic region. While shorter fragments were observed for the non-primate (Cv: *Cynocephalus variegatus* and Tbe: *Tupaia belangeri*) and strepsirrhine representatives (Lc: *Lemur catta*, Lt: *Loris tardigradus*), the tarsiers (Tb: *Tarsius bancanus*, Ts: *Tarsius syrichta*) and all anthropoid primates (Hs: *Homo sapiens*, Mm: *Macaca mulatta*, Cg: *Colobus guereza*, Aa: *Aotus azarae* and Sf: *Saguinus fuscicollis*) exhibit fragments that are about 150 bp larger in size. The most parsimonious explanation is represented in Fig. 4B and schematically displays the integration of the FLAM (designated Alu C14) in the lineage to the most recent common ancestor of tarsiers and anthropoid primates after the strepsirrhine split-off.

Together with the SINE integrations previously described (Alu C7, C9 and C12; Schmitz et al., 2001) this represents four strong molecular cladistic arguments in favor of a sister group relationship of tarsiers and anthropoid primates. Interestingly, the very SINE is identified as a FLAM A sequence supposed to constitute the fusion partner of the primate specific dimeric Alu that was retropositionally active before the primate divergence. Two explanations are possible. Firstly, a deletion could have taken place eliminating exactly one half of the transposed Alu after its integration. This would have to have taken place before the tarsiers split off from the lineage leading to the anthropoids. It is proposed that Alu SINEs or their constituting monomers become completely transpositionally inactive after reintegration into the genome. However, there is growing evidence that this "dead on arrival" mechanism is not unvaryingly applicable for each SINE (see for example Skryabin et al., 1998).

The likelihood of lineage sorting confounding the interpretations based on four SINE integrations is exceedingly low. Taken together, mono-locus evidence described here and in Schmitz et al. (2001) as well as multi-locus evidence from different SINEs (Zietkiewicz et al., 1999; Kuryshev et al., 2001), the monophyly of haplorrhine primates (Anthropoidea and Tarsiiformes) is strengthened by an entirely new perspective.

For the catarrhines (Cercopithecidae and Hominoidea) and platyrrhines (New World monkeys), together constituting the group of higher or anthropoid primates, other phylogenetic questions remain a focal point of debate.

Platyrrhines, though inhabiting only one landmass, are enormously diverse in terms of anatomy, behavior, mating strategies, social systems, locomotion and feeding adaptations. Altogether, 16 New World monkey genera exist. The number of species united in these genera however fluctuates since the accepted taxonomy is now considered overly split or too strongly lumped.

New World monkey phylogeny research is hampered by the fact that the extant platyrrhine diversity probably results from a radiation that took place after the platyrrhine ancestors entered South-central America. In addition, the scant fossil record renders it extremely difficult to define primitive features as a prerequisite for the determination of synapomorphies. Morphology-based phylogenetic research on platyrrhines is therefore partly contradictory and the subject of controversy.

The most basic issue is that of New World monkey origin and the question of how the platyrrhine ancestors colonized their present geographic range. Several scenarios exist for platyrrhine origins and transocean migration events. Although all of these require a crossing of water barriers or lowering of sea levels, several features such as skin histology, hair follicle arrangement, immunological and cytogenetic findings appear to reject a common origin of extant platyrrhine taxa. This would postulate repeated transocean migration.

SINE data from three independent loci strongly point to a monophyletic origin of extant New World monkeys (Singer et al., 2003) corroborating SINE evidence in the Major Histocompatibility Complex (Kriener et al., 2001) and other molecular data based on sequence comparisons.

Retropositional data are also at hand for the claw-bearing Callithrichids, the "dwarfs" among the simian primates that with the exception of *Callimico* regularly give birth to twins. The phylogenetic affiliations among the callitrichine genera *Callithrix*, *Saguinus*, *Leontopithecus* and *Callimico* and their relation to the other New World monkeys have therefore long been the subject of debate.

A first SINE analysis with several platyrrhine genera generated retropositional indications supporting a monophyly of all callithrichines including *Callimico.* Findings also seemed to support a close phylogenetic association of callitrichines with *Cebus*, *Saimiri* and *Aotus* (Singer et al., 2003). No retropositional evidence has yet been gained for other sister group relationships among platyrrhines. From a SINE-based perspective at least, platyrrhine phylogenetic relationships have therefore not been satisfactorily resolved. More data are urgently required to eliminate the notorious lineage sorting problem since platyrrhine history, characterized by radiation like evolution, is predestinated to cause problems in this respect.

Catarrhines, the cercopithecoid monkeys and hominoids, represent the youngest group of primates. The main difference to the arboreal New World monkeys is that these mainly ground-dwelling animals were able to invade the grasslands, thus adapting to an entirely new ecological situation. Extant catarrhines include the Cercopithecoidea (Cercopithecinae and Colobinae) and the hominoids including the gibbons, orangutans, gorillas, chimpanzees and bonobos. With the exception of humans, their geographic distribution is restricted to Africa and Asia. Fossil records however can be interpreted to place apes and monkeys in Eurasia at 5 to 17 MYA. A synthetic view on catarrhine evolution combining neontological – including molecular – and palaeontological data is at hand (Stewart and Disotell, 1998). All types of evidence are remarkably congruent. In contrast to the relatively clear phylogenetic relationships among the Papionini, a group of catarrhines consisting of the genera *Macaca*, *Papio*, *Theropithecus*, *Lophocebus*, *Mandrillus*, *Cercocebus* (Page and Goodman, 2001), the affiliations among the Cercopithecini (comprising *Cercopithecus*, *Chlorocebus*, *Erythrocebus*, *Allenopithecus*, *Miopithecus*) is probably in greater dispute. However, many of the splitting events discussed – see e.g. the macaque radiation – will not be suitable for SINE-based analysis since the splits are usually recent or assumed to be radiation-like, increasing the likelihood for polymorphic SINEs to be encountered. One question apparently well suited to SINE-based investigation is that of monophyly of Asian and African colobines. How often colobines colonized the Asian landmass is still a matter of dispute. Molecular cladistic approaches could help to approach this problem from a new perspective. However, the unavailability of material suitable for molecular analyses renders thorough taxonomic sampling nearly impossible. SINE examinations require a certain amount of relatively undegraded DNA, which is a clear drawback compared to the fecal sample-based mtDNA analyses for instance. Sample availability is also a problem for the gibbons, a group of "small" apes. Molecular sequence data (e.g. Roos and Geissmann, 2001) and chromosomal analyses allow the gibbons to be classed into four different genera *Nomascus*, *Symphalangus*, *Bunopithecus* and *Hylobates*. Corroboration of

the exact branching pattern by SINEs however is hampered by the fact that genetic material from animals not crossbred with other members of the same genus – a problem usually encountered in zoo populations – is difficult to access. Recently, SINE evidence for the major hominoid splits was created giving retropositional evidence e.g. for a clear sister group relationship between chimpanzees and humans while clearly distinguishing both lineages (Salem et al., 2003b).

For all recent splits, SINEs will be of great value as a population genetic marker system (Batzer and Deininger, 2002). Recently transposing Alus, as is the case for the human Ya, Yb, Yc, and Yd sequences (Carroll et al., 2001; Roy-Engel et al., 2001), are likely to be encountered in other species, too. Growing data sets will lead to a situation where the real impact of SINEs on the architecture of primate genomes can finally be estimated, as already discernible in recent human evolution. Large data sets will allow the characterization of new insertion polymorphisms, lineage specific deletions and gene conversions for those Alu integrations (Salem et al., 2003a, b). Apart from small changes in regulatory regions or major chromosomal rearrangements by breakage, religation or segmental duplication, the position of interspersed repeat sequences constitute additional important forces possibly altering gene function and regulation, potentially resulting in major phenotypic changes (Samonte and Eichler, 2002).

Incoming primate sequence data will provide first indications of a possible regulatory function of at least several SINE sequences in the genomes of the different primates or other taxa.

Thus, retroposon analyses are worthwhile experimental approaches for both pattern and process in primate evolutionary research. Our recording and understanding of the molecular differences between the single primate taxa are still limited. Also, we are only just beginning to appreciate the fascinating uniqueness of individual primate taxa, quite apart from their mutual phylogenetic affiliations and affiliation to humans. Given the threatened status of a considerable number of the animals discussed above, we would do well to remember our responsibility for preserving this uniqueness and variability.

Acknowledgement

Thanks go to an anonymous reviewer for comments that improved the final version of the manuscript. We express our gratitude to K. Gee for editorial assistance.

References

Andrews TD, Jermiin LS, Easteal S: Accelerated evolution of cytochrome b in simian primates: adaptive evolution in concert with other mitochondrial proteins? J Mol Evol 47:249–257 (1998).

Arnason U, Gullberg A, Janke A: Molecular timing of primate divergences as estimated by two nonprimate calibration points. J Mol Evol 47:718–727 (1998).

Arnason U, Adegoke JA, Bodin K, Born EW, Esa YB, Gullberg A, Milsson M, Short RV, Xu XF, Janke A: Mammalian mitogenomic relationships and the root of the eutherian tree. Proc Natl Acad Sci USA 99:8151–8156 (2002).

Batzer MA, Deininger PL: A human-specific subfamily of Alu sequences. Genomics 9:481–487 (1991).

Batzer MA, Deininger PL: Alu repeats and human genomic diversity. Nat Rev Genet 3:370–379 (2002).

Batzer MA, Stoneking M, Alegria-Hartman M, Bazan H, Kass DH, Shaikh TH, Novick GE, Ioannou PA, Scheer WD, Herrera RJ, et al: African origin of human-specific polymorphic Alu insertions. Proc Natl Acad Sci USA 91:12288–12292 (1994).

Batzer MA, Deininger PL, Hellmann-Blumberg U, Jurka J, Labuda D, Rubin CM, Schmid CW, Zietkiewicz E, Zuckerkandl E: Standardized nomenclature for Alu repeats. J Mol Evol 42:3–6 (1996).

Beard KC: Gliding behaviour and palaeoecology of the alleged primate family Paromomyidae (Mammalia, Dermoptera). Nature 345:340–341 (1990).

Beard KC: Origin and evolution of gliding in early Cenozoic Dermoptera (Mammalia, Primatomorpha), in Fleagle JG, Macphee RDE (eds): Primates and Their Relatives in Phylogenetic Perspective, pp 63–90 (Plenum Press, New York 1993).

Britten RJ, Baron WF, Stout DB, Davidson EH: Sources and evolution of human Alu repeated sequences. Proc Natl Acad Sci USA 85:4770–4774 (1988).

Brosius J: RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. Gene 238:115–134 (1999).

Cantrell MA, Filanoski BJ, Ingermann AR, Olsson K, DiLuglio N, Lister Z, Wichman HA: An ancient retrovirus-like element contains hot spots for SINE insertion. Genetics 158:769–777 (2001).

Carroll ML, Roy-Engel AM, Nguyen SV, Salem AH, Vogel E, Vincent B, Myers J, Ahmad Z, Nguyen L, Sammarco M, et al: Large-scale analysis of the Alu Ya5 and Yb8 subfamilies and their contribution to human genomic diversity. J Mol Biol 311:17–40 (2001).

Cheng JF, Printz R, Callaghan T, Shuey D, Hardison RC: The rabbit C family of short, interspersed repeats. Nucleotide sequence determination and transcriptional analysis. J Mol Biol 176:1–20 (1984).

Deininger PL, Batzer MA: Evolution of retroposons. Evol Biol 27:157–196 (1993).

Deininger PL, Batzer MA, Hutchison CA 3rd, Edgell MH: Master genes in mammalian repetitive DNA amplification. Trends Genet 8:307–311 (1992).

Feng Q, Moran JV, Kazazian HH Jr, Boeke JD: Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87:905–916 (1996).

Geissmann T: Vergleichende Primatologie (Springer-Verlag, Berlin 2002).

Groves CP: Primate Taxonomy (Smithsonian Institution Press, Washington DC 2001).

Hamdi H, Nishio H, Zielinski R, Dugaiczyk A: Origin and phylogenetic distribution of Alu DNA repeats: irreversible events in the evolution of primates. J Mol Biol 289:861–871 (1999).

Johanning K, Stevenson CA, Oyeniran OO, Gozal YM, Roy-Engel AM, Jurka J, Deininger PL: Potential for retroposition by old Alu subfamilies. J Mol Evol 56:658–664 (2003).

Jurka J: Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. Proc Natl Acad Sci USA 94:1872–1877 (1997).

Jurka J, Klonowski P, Trifonov EN: Mammalian retroposons integrate at kinkable DNA sites. J Biomol Struct Dyn 15:717–721 (1998).

Kajikawa M, Okada N: LINEs mobilize SINEs in the eel through a shared 3′ sequence. Cell 111:433–444 (2002).

Kay RF, Thorington RW, Houde P: Eocene plesiadapiform shows affinities with flying lemurs not primates. Nature 345:342–344 (1990).

Kazazian HH Jr, Moran JV: The impact of L1 retrotransposons on the human genome. Nat Genet 19:19–24 (1998).

Korenberg JR, Rykowski MC: Human genome organization: Alu, lines, and the molecular structure of metaphase chromosome bands. Cell 53:391–400 (1988).

Kriener K, O'hUigin C, Klein J: Alu elements support independent origin of prosimian, platyrrhine, and catarrhine Mhc-DRB genes. Genome Res 10:634–43 (2001).

Kuryshev VY, Skryabin BV, Kremerskothen J, Jurka J, Brosius J: Birth of a gene: locus of neuronal BC200 snmRNA in three prosimians and human BC200 pseudogenes as archives of change in the Anthropoidea lineage. J Mol Biol 309:1049–1066 (2001).

Li WH, Gu Z, Wang H, Nekrutenko A: Evolutionary analyses of the human genome. Nature 409:847–849 (2001).

Luan DD, Korman MH, Jakubczak JL, Eickbush TH: Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72:595–605 (1993).

Madsen O, Scally M, Douady CJ, Kao DJ, Debry RW, Adkins R, Amrine HM, Stanhope MJ, de Jong WW, Springer MS: Parallel adaptive radiations in two major clades of placental mammals. Nature 409:610–614 (2001).

Makalowski W, Mitchell GA, Labuda D: Alu sequences in the coding regions of mRNA: a source of protein variability. Trends Genet 10:188–193 (1994).

Malik HS, Burke WD, Eickbush TH: The age and evolution of non-LTR retrotransposable elements. Mol Biol Evol 16:793–805 (1999).
Martin RD: Primate origin and evolution: a phylogenetic reconstruction (Chapman Hall, London 1990).
Martin RD: Primate origins: plugging the gaps. Nature 363:223–234 (1993).
Matera AG, Hellmann U, Hintz MF, Schmid CW: Recently transposed Alu repeats result from multiple source genes. Nucleic Acids Res 18:6019–6023 (1990).
Murata S, Takasaki N, Saitoh M, Okada N: Determination of the phylogenetic relationships among Pacific salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. Proc Natl Acad Sci USA 90:6995–6999 (1993).
Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ: Molecular phylogenetics and the origins of placental mammals. Nature 409:614–618 (2001a).
Murphy WJ, Eizirik E, O'Brien SJ, Madsen O, Scally M, Douady CJ, Teeling E, Ryder OA, Stanhope MJ, de Jong WW, Springer MS: Resolution of the early placental mammal radiation using bayesian phylogenetics. Science 294:2348–2351 (2001b).
Napier JR, Napier PH: A handbook of living primates (Academic Press, London 1967).
Nekrutenko A, Li WH: Transposable elements are found in a large number of human protein-coding genes. Trends Genet 17:619–621 (2001).
Nikaido M, Nishihara H, Hukumoto Y, Okada N: Ancient SINEs from African endemic mammals. Mol Biol Evol 20:522–527 (2003).
Nishihara H, Terai Y, Okada N: Characterization of novel Alu- and tRNA-related SINEs from the tree shrew and evolutionary implications of their origins. Mol Biol Evol 19:1964–1972 (2002).
Okada N: SINEs. Curr Opin Genet Dev 1:498–504 (1991a).
Okada N: SINEs: short interspersed repeated elements of the eucaryotic genome. Trends Ecol Evol 6:358–361 (1991b).
Page SL, Goodman M: Catarrhine phylogeny: noncoding DNA evidence for a diphyletic origin of the mangabeys and for a human-chimpanzee clade. Mol Phylogenet Evol 18:14–25 (2001).
Perna NT, Batzer MA, Deininger PL, Stoneking M: Alu insertion polymorphism: a new type of marker for human population studies. Hum Biol 64:641–648 (1992).
Piskurek O, Nikaido M, Boeadi, Baba M, Okada N: Unique mammalian tRNA-derived repetitive elements in dermopterans: The t-SINE family and its retrotransposition through multiple sources. Mol Biol Evol 20:1659–1668 (2003).
Pumo DE, Finamore PS, Franek WR, Phillips CJ, Tarzami S, Balzarano D: Complete mitochondrial genome of a Neotropical fruit bat, *Artibeus jamaicensis*, and a new hypothesis of the relationships of bats to other eutherian mammals. J Mol Evol 47:709–717 (1998).
Quentin Y: Fusion of a free left Alu monomer and a free right Alu monomer at the origin of the Alu family in the primate genomes. Nucleic Acids Res 20:487–493 (1992a).
Quentin Y: Origin of the Alu family: a family of Alu-like monomers gave birth to the left and the right arms of the Alu elements. Nucleic Acids Res 20:3397–3401 (1992b).
Quentin Y: A master sequence related to a free left Alu monomer (FLAM) at the origin of the B1 family in rodent genomes. Nucleic Acids Res 22:2222–2227 (1994).
Roos C, Geissmann T: Molecular phylogeny of the major hylobatid divisions. Mol Phylogenet Evol 19:486–494 (2001).
Roos C, Schmitz J, Zischler H: Primate jumping genes elucidate strepsirrhine phylogeny. Proc Natl Acad Sci USA 101:10650–10654 (2004).
Rowe N: The Pictorial Guide to the Living Primates (Pogonias Press, East Hampton 1996).
Roy-Engel AM, Carroll ML, Vogel E, Garber RK, Nguyen SV, Salem AH, Batzer MA, Deininger PL: Alu insertion polymorphisms for the study of human genomic diversity. Genetics 159:279–290 (2001).
Ryan SC, Dugaiczyk A: Newly arisen DNA repeats in primate phylogeny. Proc Natl Acad Sci USA 86:9360–9364 (1989).
Sakamoto K, Okada N: Rodent type 2 Alu family, rat identifier sequence, rabbit C family, and bovine or goat 73-bp repeat may have evolved from tRNA genes. J Mol Evol 22:134–140 (1985).
Salem AH, Kilroy GE, Watkins WS, Jorde LB, Batzer MA: Recently integrated Alu elements and human genomic diversity. Mol Biol Evol 20:1349–1361 (2003a).
Salem AH, Ray DA, Xing J, Callinan PA, Myers JS, Hedges DJ, Garber RK, Witherspoon DJ, Jorde LB, Batzer MA: Alu elements and hominid phylogenetics. Proc Natl Acad Sci USA 100:12787–12791 (2003b).
Samonte RV, Eichler EE: Segmental duplications and the evolution of the primate genome. Nat Rev Genet 3:65–72 (2002).
Schmid CW: Alu: structure, origin, evolution, significance and function of one-tenth of human DNA. Prog Nucleic Acid Res Mol Biol 53:283–319 (1996).
Schmid C, Maraia R: Transcriptional regulation and transpositional selection of active SINE sequences. Curr Opin Genet Dev 2: 874–882 (1992).
Schmitz J, Zischler H: A novel family of tRNA-derived SINEs in the colugo and two new retrotransposable markers separating dermopterans from primates. Mol Phylogenet Evol 28:341–349 (2003).
Schmitz J, Ohme M, Zischler H: SINE insertions in cladistic analyses and the phylogenetic affiliations of *Tarsius bancanus* to other primates. Genetics 157:777–784 (2001).
Schmitz J, Ohme M, Zischler H: The complete mitochondrial sequence of *Tarsius bancanus*: evidence for an extensive nucleotide compositional plasticity of primate mitochondrial DNA. Mol Biol Evol 19:544–553 (2002a).
Schmitz J, Ohme M, Suryobroto B, Zischler H: The colugo (*Cynocephalus variegatus*, Dermoptera): The primates' gliding sister? Mol Biol Evol 19:2308–2312 (2002b).
Shedlock AM, Okada N: SINE insertions: powerful tools for molecular systematics. Bioessays 22:148–160 (2000).
Shimamura M, Yasue H, Ohshima K, Abe H, Kato H, Kishiro T, Goto M, Munechika I, Okada N: Molecular evidence from retroposons that whales form a clade within even-toed ungulates. Nature 388:666–670 (1997).
Shimamura M, Abe H, Nikaido M, Ohshima K, Okada N: Genealogy of families of SINEs in cetaceans and artiodactyls: the presence of a huge superfamily of tRNA(Glu)-derived families of SINEs. Mol Biol Evol 16:1046–1060 (1999).
Shoshani J, Groves CP, Simons EL, Gunnell GF: Primate phylogeny: morphological vs. molecular results. Mol Phylogenet Evol 5:102–154 (1996).
Singer SS, Schmitz J, Schwiegk C, Zischler H: Molecular cladistic markers in New World monkey phylogeny (Platyrrhini, Primates). Mol Phylogenet Evol 26:490–501 (2003).
Skryabin BV, Kremerskothen J, Vassilacopoulou D, Disotell TR: The BC200 RNA gene and its neural expression are conserved in Anthropoidea (Primates). J Mol Evol 47:677–685 (1998).
Smit AF, Riggs AD: MIRs are classic, tRNA-derived SINEs that amplified before the mammalian radiation. Nucleic Acids Res 23:98–102 (1995).
Starck D: Vergleichende Anatomie der Wirbeltiere auf evolutionsbiologischer Grundlage (Springer Verlag, Berlin 1978).
Stewart CB, Disotell TR: Primate evolution – in and out of Africa. Curr Biol 13:R582–588 (1998).
Sverdlov ED: Retroviruses and primate evolution. Bioessays 22:161–71 (2000).
Takahashi K, Terai Y, Nishida M, Okada N: Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons. Mol Biol Evol 18:2057–2066 (2001).
Tavaré S, Marshall CR, Will O, Soligo C, Martin RD: Using the fossil record to estimate the age of the last common ancestor of extant primates. Nature 416:726–729 (2002).
Teeling EC, Scally M, Kao DJ, Romagnoli ML, Springer MS, Stanhope MJ: Molecular evidence regarding the origin of echolocation and flight in bats. Nature 403:188–192 (2000).
Verneau O, Catzeflis F, Furano AV: Determining and dating recent rodent speciation events by using L1 (LINE-1) retrotransposons. Proc Natl Acad Sci USA 95:11284–11289 (1998).
Yoder AD: The phylogenetic position of genus *Tarsius*: whose side are you on? in Wright PC, Simons EL, Gursky S (eds): Tarsiers: Past, Present, and Future, pp 161–175 (Rutgers University Press, New Brunswick 2003).
Zietkiewicz E, Richer C, Labuda D: Phylogenetic affinities of tarsier in the context of primate Alu repeats. Mol Phylogenet Evol 11:77–83 (1999).