J. Zool. Syst. Evol. Research **38** (2000) 65–72
© 2000 Blackwell Wissenschafts-Verlag, Berlin
ISSN 0947-5745

Received on 31 December 1999

Institute for Zoo Biology and Wildlife Research, Berlin, Germany

Use of mitochondrial DNA sequences to test the Ceratomorpha (Perissodactyla:Mammalia) hypothesis

C. PITRA and J. VEITS

Abstract

Wood (J. Mamm. 18, 106–118; 1937) united the superfamilies Tapiroidea *sensu stricto* and Rhinocerotoidea in the suborder Ceratomorpha and aligned the Ceratomorpha with the suborder Hippomorpha within the order Perissodactyla. Although the monophyly of the Ceratomorpha appears now well-supported in paleontological and morphological analyses, the molecular relationship among the three extant superfamilies Tapiroidea, Rhinocerotoidea, and Equoidea has not yet been examined due to the limited amount of molecular information on tapirs. In the present study, we examined the phylogenetic position of Tapiroidea, represented by the complete mitochondrial cytochrome *b* gene (1140 bp) of a lowland tapir (*Tapirus terrestris*), and a Indian tapir (*Tapirus indicus*), relative to modern horses, zebras, donkeys, and rhinoceroses. The phylogenetic analyses using standard parsimony, neighbour-joining and maximum likelihood algorithms revealed monophyly of the Perissodactyla and three clearly distinct lineages: the modern horses, tapirs, and rhinoceroses. However, the sister-taxon relationship of the tapirs to either the rhinoceroses or the horses was not resolved conclusively in the bootstrap analysis. Spectral analysis, in which phylogenetic information is displayed independently of any selected tree, revealed that the DNA sequences available do not contain enough phylogenetic signal for any of the alternative hypotheses on the basal diversification of perissodactyls. The short branch lengths among the three perissodactyl lineages suggest that they diverged within a relatively short period, a finding consistent with molecular divergence datings and the fossil evidence that indicates a major radiation of the early perissodactyls approximately 54–50 million years ago.

Key words: Perissodactyla – Ceratomorpha – tapirs – molecular taxonomy – cytochrome *b*

Introduction

Traditionally the extant Perissodactyla are classified into the suborder Hippomorpha, which includes the horses, asses and zebras (superfamily Equoidea) and the suborder Ceratomorpha, which includes the rhinoceroses (superfamily Rhinocerotoidea) and the tapirs (superfamily Tapiroidea) (for discussion and references see Prothero and Schoch 1989). A major suggestion is, that in the preceding scheme, the tapiroids and rhinocerotoids are regarded as being more closely related to each other than to the hippomorphs. The hypothesis of Ceratomorpha monophyly, originally proposed by Wood (1937), has been increasingly supported by several previously recognized dental and nondental osteological characters, and a few features of the soft anatomy shared by early tapiroids and primitive rhinocerotids (Radinsky 1967; Prothero et al. 1986). However, two early biomolecular studies, one on restriction site mapping of the alpha globin gene cluster (Flint et al. 1990) and the other on amino acid sequence data of pancreatic polypeptide (Henry et al. 1993), did not recognize tapirs and rhinoceroses as sister-groups opposing the Equidae. Recent molecular studies address the Perissodactyla phylogeny only in part (Xu et al. 1996; Holmes and Ellis 1999) leaving the position of the tapir lineage within the Perissodactyla unresolved.

In this paper, we present new and complete sequences of the mitochondrial cytochrome *b* (cyt *b*) gene for the Brazilian or lowland tapir (*Tapirus terrestris*), and the Indian tapir (*Tapirus indicus*), which represent the tapirid clade within the perissodactyl phylogenetic tree. On the basis of this extended data set and by including existing information from the cytochrome *c* oxidase subunit II (COII) gene (Ashley et al. 1996) we examined the molecular relationship among the Tapiroidea, Rhinocerotoidea, and Equoidea. The analysis allows further testing of Wood's (1937) hypothesis of Ceratomorpha monophyly.

We employed both standard methods of phylogenetic reconstruction as well as spectral analysis (Hendy et al. 1994) to identify the strength of the conserved phylogenetic signals. Inclusion of representatives of the Artiodactyla (cattle) and Cetacea (blue whale) permitted the application of a time-frame based on the artiodactyla/cetacean split at 60 million years ago (mya) (Arnason and Gullberg 1996) as a calibration point dating the deep perissodactyl divergences.

Materials and methods

Cytochrome *b* amplification and sequencing

Total DNA was isolated using the QIAamp blood and tissue kit (Qiagen Inc., Germany) according to manufacturer's instructions. The cyt *b* sequences were obtained from two independently amplified polymerase chain reaction (PCR) products using universal vertebrate primers: CB1-L, CB2-H (Kocher et al. 1989), CB4a-L, CB3-H, and GLUDG-L (Palumbi 1991). The PCR profile was as follows: each cycle consisting of denaturation for 0.5 min at 93°C, annealing for 1 min at 50°C, and final extension for 1 min at 72°C for 30–35 cycles. The PCR products were purified (QIAquick PCR purification kit; Qiagen), and sequenced by the extension-dideoxy-chain termination method (Sanger et al. 1977) with a commercial kit (Dye Terminator Cycle Sequencing Kit, Applied Biosystems, Foster City, CA) and analysed on a 373-A sequencer (Applied Biosystems). DNA was sequenced bidirectionally. Sequences corresponding to positions 14514–15314 in the cattle cyt *b* sequence (Anderson et al. 1982) were aligned by the CLUSTAL W algorithm (Thompson et al. 1994) and deposited in the GenBank database.

Data analyses

We computed basic sequence statistics and Kimura's (1980) two-parameter distances (K2P) with MEGA (Kumar et al. 1993). SPECTRUM 2.0 (Charleston 1996) was used to compute log-determinant distances (log-det). The phylogenetic analyses were performed both with nucleotide (nt) and amino acid (aa) sequences of two mitochondrial protein-coding genes (cyt *b* and COII). Sequences were analysed using the PHYLIP (Felsenstein 1993), and PUZZLE (Strimmer and von Haeseler 1996) packages. For the neighbour-joining (NJ) and the maximum parsimony (MP) analyses, confidence values were established by bootstrapping (Felsenstein 1985). In the maximum likelihood (ML) analysis, corresponding reliability values were established by quartet puzzling (QP) (Strimmer and von Haeseler 1996). The alternative hypotheses were tested both by spectral analysis (Hendy et al. 1994) and Kishino

and Hasegawa (1989) maximum likelihood test. To assess the molecular-rate heterogeneity, we compared the log-likelihoods of maximum likelihood trees computed with and without the application of a molecular clock (PUZZLE routines). The species names, abbreviations, and references of the formerly published sequences used in the analyses are as follows: horse (*Equus caballus*, Eca) (Xu and Arnason 1994), donkey (*Equus asinus*, Eas) (Xu et al. 1996), Grevy's zebra (*Equus grevyi*, Egr) (Irwin et al. 1991), white rhino (*Ceratotherium simum*, Csi) (Xu and Arnason 1997), Indian rhino (*Rhinoceros unicornis*, Run) (Xu et al. 1996), black rhino (*Diceros bicornis*, Dbi) (Irwin et al. 1991), lowland tapir (*Tapirus terrestris*, Tte) (Ashley et al. 1996), Indian tapir (*Tapirus indicus*, Tin) (Ashley et al. 1996), blue whale (*Balaenoptera musculus*, Bmu) (Arnason et al. 1993), and cattle (*Bos taurus*, Bta) (Anderson et al. 1982).

Results

Alignment of the tapir cytochrome *b* sequences

The newly determined sequences of the cyt *b* gene in *Tapirus terrestris* (AF056030) and *Tapirus indicus* (AF145734) are 1140 nt in length and include a deduced translation product of 379 aa in length. Percentage base composition was biased in a manner typical for other mammals where guanine residues are uncommon, especially at third codon positions (Irwin et al. 1991). Overall base percentage was: adenine (30.7%), thymine (27.6%), cytosine (29.2%), and guanine (12.5%).

Alignment of cyt *b* sequences of the tapirs and six other perissodactyls shows no nucleotide insertions or deletions within sequences. The gene sequences can be fully translated using the bovine mitochondrial code (Anderson et al. 1982) without non-sense or intervening stop codons. Finally, the alignment revealed neither an increased abundance of first and second codon position changes, nor a shift in the typical mammalian mtDNA transition bias, known to occur in mtDNA sequences translocated to the nuclear genome (Zhang and Hewitt 1996). Taken together, we consider the tapir cyt *b* sequences reported, were of solely mitochondrial origin.

There are 371 variable nucleotide positions among the eight gene sequences compared (32.6% of the total cyt *b* gene sequence) and 229 (20%) were parsimony informative. Variable nucleotides were predominantly at third codon positions (278 sites, 73.4%), followed by first positions (69 sites, 18.2%), and second positions (24 sites, 6.3%). The majority of substitutions occurred at a mean number of 281.9 ± 3.1 synonymous sites (third codon positions and leucine codon positions).

Pattern of pairwise sequence divergence and saturation

Within the Perissodactyla, the K2P-corrected third codon position distances ranged from 18.4% (donkey versus zebra) to 64.6% (lowland tapir versus horse). From Perissodactyla to the outgroup species (cattle and blue whale), the computed mean distance value is $64.4 \pm 6.1\%$ ($n = 16$) identifying the outgroup comparisons at third codon positions as being highly affected by multiple substitutions, i.e. with a zone of saturation. The log-det correction of the sequence data (instead of the K2P correction) was performed to avoid the misleading effect of unequal nucleotide composition (Lockhart et al. 1994). A pairwise comparison of the numbers of substitutions between the cyt *b* sequences (Fig. 1) partitioned into transitions (Ti) and transversions (Tv) revealed a differing pattern of divergence dependent upon codon position. The Tv differences at the third codon position between diverging taxa increases linearly, whereas the Ti changes increase linearily up to approximately 11% total difference and then enter saturation. In contrast to third codon positions, the growth of each substitution class at codon $1^\circ + 2^\circ$ positions increases linearly between all taxa in this study without evidence of a plateau in divergence. Saturation was not evident, however, when the numbers of substitutions at both nonsynonymous and synonymous sites were plotted over the range of divergences for all pairs of ingroup and outgroup taxa (Fig. 2).

Cytochrome *b* protein sequence variation

Among the 379 aa sites of the translated cyt *b* protein sequences, replacement substitutions were inferred for 84 positions (22.2%), but only 42 replacements (11.1%) were phylogenetically informative. The majority of aa replacements involved exchanges between hydrophobic residues.

Phylogenetic analyses

The DNA sequence data for 10 selected taxa, including eight perissodactyl species, was arranged in 12 ways. These arrangements were done with both the nt and aa sequences of the complete cyt *b* gene alone as well as in combination with partial sequences from the COII gene. At the nt level the analysis was based on substitutions at all codon positions, first and second positions together, and third position alone, in order to examine the degree of support provided by each character set for a particular topology. Additionally, conservative nt differences, consisting of all substitutions in first codon position (except synonymous leucine transitions), all substitutions in position 2 and transversions only in position 3 (Irwin et al. 1991), as well as all nonsynonymous substitutions were analysed separately. Three phylogenetic inference algorithms (NJ, MP, and ML) and subsequent bootstrap analyses were used to choose optimal trees. These 36 reconstructions yielded only three different trees (I, II, III) shown in Fig. 3. All three optimal consensus trees highly support a monophyletic origin for the Perissodactyla (>94% bootstrap and QP values, respectively). The order is composed of three clearly distinct clades: the modern horses, tapirs and rhinoceroses. The support for these groupings is consistently high in all methods of reconstruction (>94% bootstrap and QP values, respectively), although the topology within clades differed slightly among trees. The distinguishing feature among the Trees I, II and III is the phylogenetic relationship of the tapirs. In the various analyses of the 12 data sets, the tapir lineage grouped with either the rhinoceroses (Tree I) or the horse lineage (Tree II). It is notable that each tree building method for the aa sequences of the cyt *b* and the combined COII + cyt *b* data set yielded a tree of Type II with a highly supported tapir/horse grouping. This consistency is opposed by different trees inferred by each tree building method for the various nt sequence data sets. Although the tapir lineage now generally appears to group with the rhinoceros lineage (Tree I in Fig. 3), the bootstrap support for this internal node ranged from 33 to 100%, depending on the underlying data set. MP analysis based on nonsynonymous changes failed to resolve a higher taxon relationship and the resulting phylogenies represent the three perissodactyl families in a trichotomy (Tree III in Fig. 3). Furthermore, no arrangement, linking the rhinoceroses and the horses, is found in any of the analyses of the sequence data sets. Taken together, the intra-ordinal position of the tapirs remains the central difference between the conflicting hypotheses produced from our analyses.

Spectral analysis

To further assess how probable the various phylogenetic hypotheses are, we measured the definitive phylogenetic signals

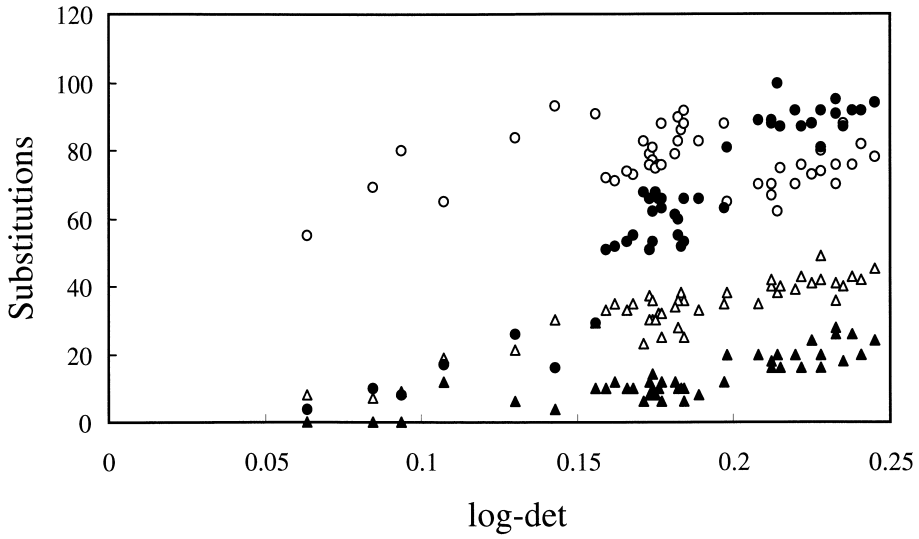

Fig. 1. Plot of the numbers of transitions and transversions at 1° + 2° and 3° codon positions of the mitochondrial cytochrome *b* gene versus log-det. Open circles are 3°-position transitions, closed circles are 3°-position transversions, open triangles are 1° + 2°-position transitions, and closed triangles are 1° + 2°-position transversions

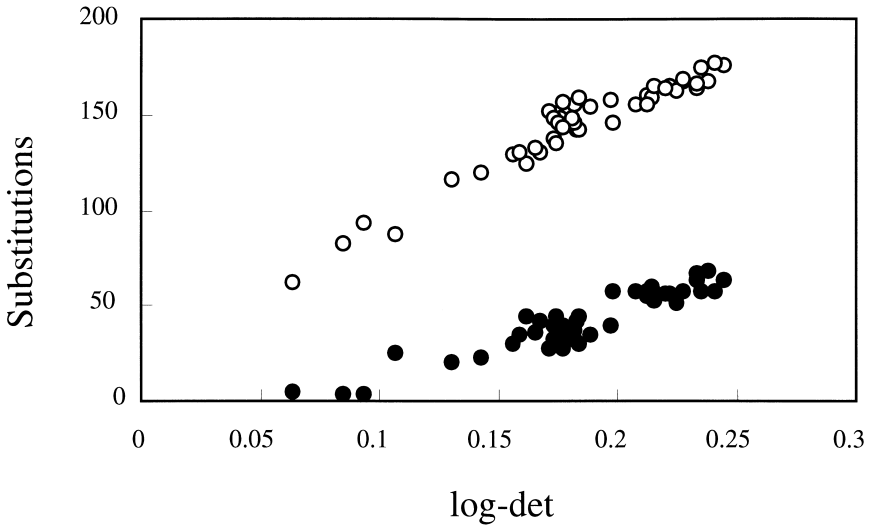

Fig. 2. Plot of nucleotide substitutions at both synonymous (open circles) and nonsynonymous sites (closed circles) of the mitochondrial cytochrome *b* gene versus log-det

in a data set using spectral analysis (Hendy et al. 1994; Lento et al. 1995). Spectral analysis requires not merely examining an optimal tree computed by a tree reconstruction method but using the original DNA sequence data themselves to look for direct support for every taxon grouping (split) within a data set. Therefore, this method can quantify the amount of support for (and conflicts against) any given phylogenetic arrangement and is extremely useful when multiple hypotheses are to be compared. Figure 4 is a histogram-like spectrum generated from the conservative nt positions of the cyt *b* (according to Irwin et al. 1991) with the two outgroups, which displays all phylogenetic information contained in that data set. With 10 taxa, there are 2^{10-1}, or 512 possible splits, most of which are equivalent to random combinations of taxa. Of these, 33 splits are supported by some evidence and these are the bars shown in the top half of the spectrum. Bars in the lower half of the spectrum show the frequency of conflicting evidence for the corresponding split. With decreasing support/conflict ratio, groupings of taxa becomes less likely. The grey bars represent singletons, or terminal taxa, and the black bars indicate those taxonomic groupings (splits) that have been included in the phylogeny identical to that presented in inset Fig. 4. This tree is unrooted and of Type I in Fig. 3. The spectrum indicates that the strongest signals (i.e. highest ratios of support to conflict) are for the internal nodes (i.e. splits 127, 112, and 12) reflecting the family structure of the Perissodactyla. Split 124, representing tapirs/rhinoceroses monophyly has slightly higher signals than split 115, which represents tapirs/horses monophyly. Although the latter has been excluded from the optimal tree (inset Fig. 4), the generally low support/conflict ratios of both splits demonstrate that there is no strong support in the data for any of the alternative phylogenetic hypotheses.

The results of the spectral analyses generated from different nt sequence data sets are presented in Fig. 5. The strength of support for specific splits in the spectral data was calculated as the frequency of support minus the frequency of conflict (*S–C*). Generally, the (*S–C*) values prove to be a better predictor than either the *S* or *C* values separately (Lento et al. 1995). The results are interesting because there are considerable differences in (*S–C*) values between the internal nodes of the tree. A high

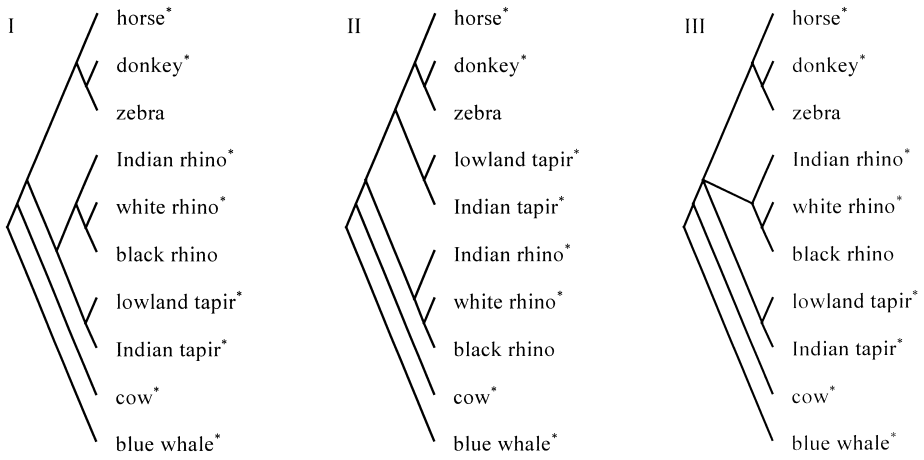

Fig. 3. The 36 phylogenetic reconstructions result in these three trees. The table below indicates which analysis results in which tree topology. The bootstrap and quartet puzzling support values, respectively, for the different sister-group positions of the tapirs to several members of the order Perissodactyla are given in parentheses. Note that values lower than 70 tend to collapse into a basal trifurcation of the type III. The asterisk (*) denote the taxa that were used in the analyses based on the sequences of the complete cytochrome *b* (cyt *b*) gene combined with partial sequences from the cytochrome *c* oxidase (COII) gene

Analysis	Sequence length (bp)	Parsimony	Neighbour-joining	Maximum-likelihood
cyt *b* (all pos)	1137	I (81)	I (47)	I (62)
cyt *b* (1°+2°)	758	II (83)	II (92)	II (91)
cyt *b* (3°)	379	I (51)	I (39)	I (82)
cyt *b* (conserv)	729	I (75)	I (65)	I (66)
cyt *b* (nonsyn)	252	I (66)	II (51)	I (65)
cyt *b* (aa)	379	II (92)	II (91)	II (97)
COII+cyt *b* (all pos)	1788	I (88)	I (76)	I (94)
COII+cyt *b* (1°+2°)	1192	II (81)	I (90)	II (95)
COII+cyt *b* (3°)	596	I (92)	I (83)	I (100)
COII+cyt *b* (conserv)	1144	I (59)	I (71)	I (66)
COII+cyt *b* (nonsyn)	291	I (70)	II (72)	II.(53)
COII+cyt *b* (aa)	596	II (96)	II (85)	II (100)

level of (*S–C*) values is given for the internal nodes separating the modern horses, rhinoceroses, and tapirs. As for the alternative groupings of tapirs with either the horses or the rhinoceroses there are consistently low (*S–C*) values in all alignments. These latter two nodes are deep in the interior of the tree where the conflict probability can be higher (Wägele and Rödding 1998). This assumption, however, is obviously not true for the cyt *b* and COII sequences in our data set because for the deepest node in the tree which separates all perissodactyls from the outgroup taxa, a clear phylogenetic signal is conserved.

Nuclear DNA sequence evidence

With respect to the labile sister-relationship of the tapirs, it is noteworthy that similar results were obtained by analyses of all nuclear gene data available at the moment (Table 1). For example, the class I and class II genes of the major histocompatibility complex (MHC) did not resolve a tapir–rhinoceros–horse triad, whereas a tapir–horse relationship was best supported by ML analysis of the epsilon–globin gene sequences. In the case of the luteinizing hormone beta-subunit (LH-beta) gene sequences, a sister-taxon relationship between the tapirs and the rhinoceroses was favoured on the basis of log-likelihood comparisons. The frequencies of support for splits of interest in the spectrum were in none of these cases higher than the frequencies of conflict, indicating that the analyses of the nuclear DNA sequences available also failed to give a clear resolution (Table 1).

Dating of divergences

The maximum likelihood distances from analyses with and without the application of a molecular clock (data not shown) did not differ significantly ($p > 0.1$), indicating molecular-rate constancy among taxa, therefore estimates of divergence times should be possible if calibration points are accurate. As an external standard, we used the calibrated artiodactyla/cetacean divergence at 60 mya (Aranson and Gullberg 1996). Maximum likelihood branch lengths were computed from a completely resolved quartet puzzling tree (1000 puzzling steps; $\log L$: -2952.2) using the HKY model of substitution (Hasegawa et al. 1985). The datings are based on the conservative nt positions of the cyt *b* and the homogenous-rate model. The cyt *b* analysis dated the divergence between the equoids and the tapiroids to 53.9 ± 9.0 mya, that between tapiroids and rhinocerotids to 51.6 ± 8.1 mya, and that between equoids and rhinocerotids to 56.9 ± 12.7 mya. The estimates were not significantly different in the analysis of variance results ($p = 0.1819$), thus we suggest that the divergence of the three lineages took place within narrow time limits and near the base of the perissodactyl evolution 54.5 mya (95% confidence limits: 52.1–57.0 mya). The evolutionary separation of the lineages leading to the modern tapirs, rhinoceroses and horses was dated to 27.0 ± 6.1, 33.7 ± 6.3,

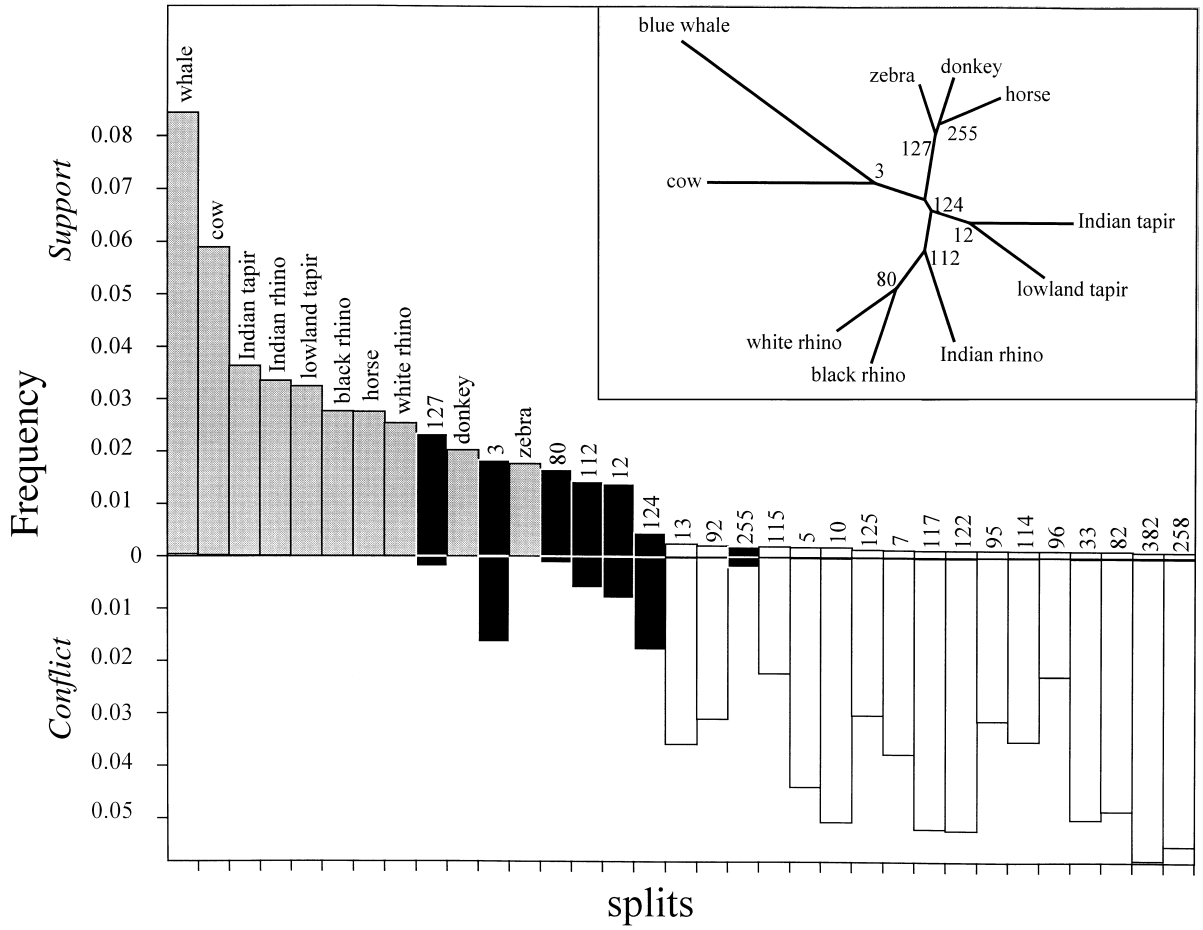

Fig. 4. Support/conflict spectrum for the conservative nucleotide positions of mitochondrial cytochrome *b* sequences from eight perissodactyls and two outgroup taxa. Bars above the *X*-axis represent frequency of support for each split. Bars below the *X*-axis represent the sum of all conflicts against the corresponding split above the *X*-axis. Solid bars (both support and conflict) indicate the splits that have been selected for inclusion in the optimal tree shown in the inset to panel 4. A number of splits of interest are marked by letters in this figure and described in the text

and 17.9 ± 5.8 mya, respectively. These datings are quite compatible with the known palaeontological record (Prothero and Schoch 1989) and the few existing molecular datings (Ashley et al. 1996; Xu and Arnason 1997) of recent perissodactyl divergences.

Discussion

The complete sequences of the mitochondrial cyt *b* gene in two recent tapirs were analysed aiming at the reconstruction of the evolutionary relationships among the Tapiroidea, Rhinocerotoidea and Equoidea. Although the cyt *b* sequences of the three superfamilies are clearcut and the monophyly of the perissodactyl sequences as a whole received strong bootstrap support, the order of divergence among the perissodactyl superfamilies remains equivocal. On the basis of the short branches among the three perissodactyl lineages, some of the resultant phylogenies in Fig. 3 were not resolved to the point of full bifurcation. Thus, it is no surprise that the exact branching pattern of the trees depend upon which sort of data matrix and model of evolution was employed during the phylogenetic analyses. Several factors may account for such inconsistencies among analyses. The short branches observed within most phylogenies may imply that the cyt *b* genes either diverged in a relatively short period of time after the origin of the Perissodactyla or that they evolved slowly in order to accumulate sufficient phylogenetic information. However, the latter possibility seems unlikely because (1) phylogenetic resolution occurrs at both basal and terminal nodes (Fig. 5); and (2) the alternative intraordinal relationships received likewise poor support in analyses both in combined mitochondrial genes (Fig. 3) and in different nuclear genes (Table 1). Therefore, the most conservative interpretation of the molecular evidence is to favour a contemporaneous origin of the three perissodactyl lineages, a so-called 'star phylogeny'. In other words, the molecular trees indicate uncertainty about the position of the tapirs, but that does not necessarily mean that a basal three-way split of a common ancestral species is the best explanation of the data. The limited phylogenetic signal of the internal nodes shown by both bootstrap analysis (Fig. 3) and spectral analyses (Fig. 5 and Table 1) could be the result of a rapid radiation after which the synonymous positions have become saturated with nucleotide substitutions, and nonsynonymous changes have not accumulated in sufficient numbers for meaningful phylogenetic analyses. This situation, which was also described in cetaceans (Arnason and Gullberg 1996), leporids (Halanych and Robinson 1999), and bovids (Pitra et al. 1998) is likely to occur in

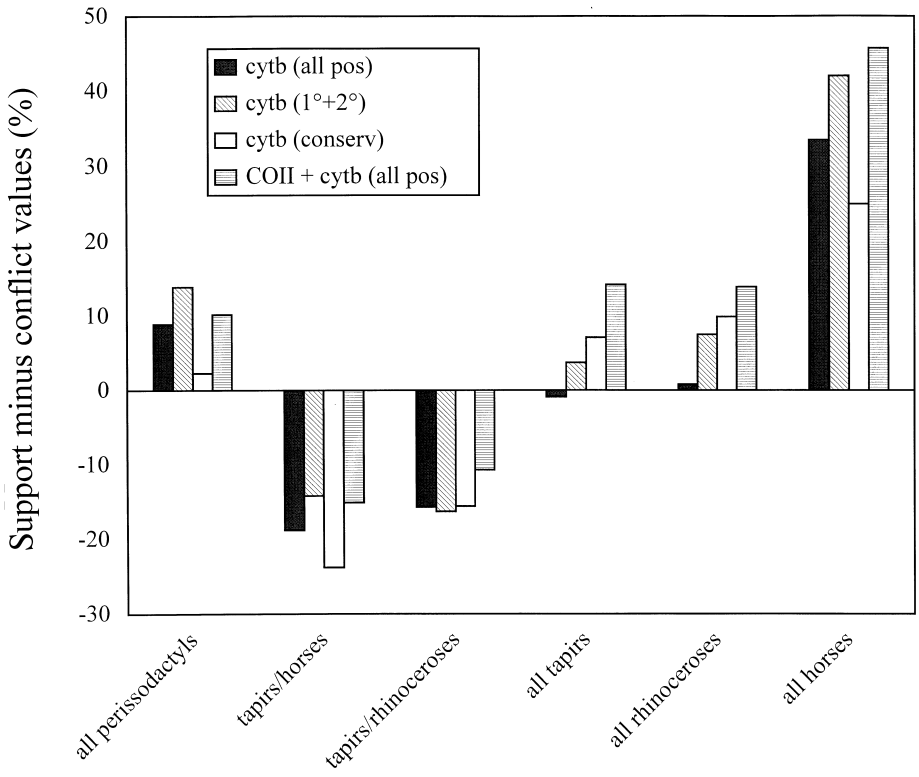

Fig. 5. Comparison between support (*S*) minus conflict (*C*) values for specific splits in the spectral data generated from different sequence data sets. The support minus conflict values (*S*–*C*) are reported as a percentage of the highest (*S*–*C*) value found in the corresponding spectrum to give support and conflict equal contribution to the overall assessment of each split

Table 1. Maximum-likelihood and spectral analysis of Perissodactyla relationship based on nucleotide sequences of several nuclear genes in rhinoceroses (RHI), tapirs (TAP) and horses (HOR). The orthologous sequence of the artiodactyl *Bos taurus* (ART) has been used as outgroup

Gene	Length (bp)	Maximum-likelihood analysis		Spectral analysis (support–conflict values)	
		Best tree	Log *L*	TAP/RHI split	TAP/HOR split
MHC class I[1]	206	(ART (RHI, TAP, HOR))	−621.56	−0.0526	−0.0367
MHC class II[2]	180	(ART (RHI, TAP, HOR))	−451.78	−0.0173	−0.0324
Epsilon-Globin[3]	1076	(ART (RHI, (TAP, HOR)))	−2598.51	−0.0028	−0.0018
LH-beat[4]	172	(ART (HOR, (TAP, RHI)))	−434.95	−0.0028	−0.0479

[1] sequences from Holmes and Ellis 1999.
[2] sequences from Fraser & Ryder, unpublished.
[3] sequences from Kim et al. unpublished.
[4] sequences from Fischer & Veits, unpublished.

other genes, such as the mitochondrial COII and nuclear genes.

The Ceratomorpha hypothesis regarding the relationship of the tapiroids to the rhinocerotoids has rested primarily on morphological characters and a reasonably abundant fossil record. Like the mitochondrial sequence data, the palaeontological data suggest that these mammals experienced a rapid diversification. On the basis of palaeontological evidence, the phylogeny of the Tapiroidea can be traced back to the early Eocene tapiroid *Homogalax protapirinus* (Schoch 1989). The oldest known rhinocerotid is *Hyrachyus*, which first appears in the late early Eocene (Prothero et al. 1986). The hippomorph radiation began in the latest Paleocene and earliest Eocene with the 'wastebasket' taxon *Hyracotherium* (Prothero and Schoch 1989). Assuming that the fossil remains are correctly placed within their phylogenetic context, the ancestor(s) of the three perissodactyl lineages obviously diversified over a relatively short time-period, from approximately 54–48 million years ago.

Previous results on perissodactyls have shown that morphological differentiation may also occur very rapidly in conjunction with adaptation to a new ecological niche (Radinsky 1969). The Perissodactyla arose from their close relatives, the tethytheres (proboscideans, sirenians) and the arsinoitheres in the late Paleocene. That process was marked by a major adaptive shift involving two functional complexes, the specialization of the limbs for running, and a change in molar cusp pattern and jaw musculature that increased transverse shear (Prothero and Schoch 1989). An initial adaptive radiation occurred shortly after the origin of the Perissodactyla producing slight differences in molar cusp pattern and size, which indicate differences in masticatory action between the earliest equoids (*Hyracotherium*) and tapiroids (*Homoglax*). Secondary radiations followed soon thereafter and the Rhinocerotoidea appeared at the beginning of the late Eocene, apparently derived from the tapiroid lineage (Radinsky 1969). The main adaptive changes

involved the molar teeth that differ from those of tapiroids by having a higher crown.

Although the molecular data (such as mtDNA) are not sufficient to challenge the traditional view of Ceratomorpha which is based on morphological characters, it is certainly sufficient to identify points for further inquiry. The lability of the molecular relationship among the extant perissodactyl lineages suggests that any proposal for sister-group relationship of the tapirs should be expressed with caution.

Since the preparation of this paper a phylogenetic investigation of the Perissodactyla using two mitochondrial genes, the cytochrome *c* oxidase subunit II gene and a portion of the 12SrRNA gene has been published (Norman and Ashley 2000). The support for Ceratomorpha is weak in this analysis (57–71% bootstrap values).

Acknowledgements

We thank J. Fickel and an anonymous reviewer for critical and beneficial advice on the first version, and S. Auls for expert technical help.

Zusammenfassung

Verwendung von mitochondrialen DNA Sequenzen zur Prüfung der Ceratomorpha (*Perissodactyla*:*Mammalia*) *Hypothese*

Wood (J. Mamm. 18, 106–118; 1937) faßte die Superfamilien Tapiroidea *sensu stricto and* Rhinocerotoidea in der Subordung Ceratomorpha zusammen und vereinigte die Subordnung Ceratomorpha mit der Subordnung Hippomorpha in der Ordnung Perissodactyla. Obwohl die Monophylie der Ceratomorpha in paläontologischen und morphologischen Analysen gut unterstützt ist, wurden die molekularen Beziehungen zwischen den rezenten Superfamilien Tapiroidea, Rhinocerotoidea und Equoidea aufgrund fehlender molekularer Information über die Tapire bisher nicht untersucht. In der vorliegenden Studie untersuchen wir die phylogenetische Stellung der Tapiroidea, repräsentiert durch das vollständige mitochondriale Cytochrom b Gen (1140 bp) eines Flachlandtapirs (*Tapirus terrestris*) und eines Schabrackentapirs (*Tapirus indicus*) relativ zu rezenten Pferden, Zebras, Eseln und Nashörnern. Die phylogenetischen Analysen unter Anwendung von Parsimonie-, Neighbor-Joining-und Maximum Likelihood-Algorithmen ergaben eine Monophylie der Perissodactyla und drei klar getrennte Linien – die modernen Pferde, Tapire und Nashörner. Die Geschwister-Taxon Beziehung der Tapire zu den Nashörnern oder den Pferden, wurde aber in der bootstrap Analyse nicht überzeugend aufgelöst. Die Spektral Analyse bei der die phylogenetische Information unabhängig von einem gewählten Stammbaum angezeigt wird, ergab, daß die verfügbaren DNA Sequenzen nicht genügend phylogenetische Information für eine der alternativen Hypothesen über die basale Diversifikation der Unpaarhufer enthalten. Die kurzen Astlängen zwischen den drei Unpaarhuferlinien legen nahe, daß sie in einer relativ kurzen Zeitperiode divergiert sind, ein Befund, der mit molekularen Divergenzdatierungen und den fossilen Belegen übereinstimmt, die eine größere Radiation der frühen Perissodactyla vor etwa 54–50 Millionen Jahren anzeigen.

References

Anderson, S.; de Bruijn, M. H.; Coulson, A. R.; Eperon, I. C.; Sanger, F.; Young, I. G., 1982: Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. J. Mol. Biol. **156,** 683–717.

Arnason, U.; Gullberg, A., 1996: Cytochrome *b* nucleotide sequences and the identification of five primary lineages of extant cetaceans. Mol. Biol. Evol. **13,** 407–417.

Arnason, U.; Gullberg, A.; Widegren, B., 1993: Cetacean mitochondrial DNA control region: sequences of all extant baleen whales and two sperm whale species. Mol. Biol. Evol. **10,** 960–970.

Ashley, M. V.; Norman, J. E.; Stross, L., 1996: Phylogenetic analysis of the perissodactylan family Tapiridae using cytochrome c oxidase (COII) sequences. J. Mammal. Evol. **3,** 315–326.

Charleston, M. A., 1996: 'Spectrum', Version 1.0.5. Glasgow, Scotland: University of Glasgow.

Felsenstein. J., 1985: Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39,** 783–791.

Felsenstein, J., 1993: PHYLIP-(Phylogeny Inference Package) 3.5c. Distributed by the author. Seattle: University of Washington,

Flint, J.; Ryder, O. A.; Clegg, J. B., 1990: Comparison of the α-globin gene cluster structure in Perissodactyla. J. Mol. Biol. **30,** 36–42.

Halanych, K. M.; Robinson, T. J., 1999: Multiple substitutions affect the phylogenetic utility of cytochrome *b* and 12S rDNA data: examining a rapid radiation in leporid. J. Mol. Evol. **48,** 369–379.

Hasegawa, M.; Kishino, H.; Yano, K., 1985: Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. **22,** 160–174.

Hendy, M. D.; Steel, M. A.; Penny, D., 1994: A discrete Fourier analysis for evolutionary trees. Proc. Natl. Acad. Sci. USA **91,** 3339–3343.

Henry, J. S.; Lance, V. A.; Conlon, J. M., 1993: Purification and characterization of insulin and the C-peptide of proinsulin from Przewalski's horse, zebra, rhino, and tapir (Perissodactyla). Gen. Comp. Endocrin. **89,** 299–308.

Holmes, E. C.; Ellis, S. A., 1999: Evolutionary history of MHC class I genes in the mammalian order Perissodactyla. J. Mol. Evol. **49,** 316–324.

Irwin, D. M.; Kocher, T. D.; Wilson, A. C., 1991: Evolution of the cytochrome *b* gene of mammals. J. Mol. Evol. **32,** 128–144.

Kimura, M., 1980: A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16,** 111–120.

Kishino, H.; Hasegawa, M., 1989: Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and branching order in Hominoidea. J. Mol. Evol. **29,** 170–179.

Kocher, T. D.; Thomas, W. K.; Meyer, A.; Edwards, S. V.; Pääbo, S.; Villablanca, F. X.; Wilson, A. C., 1989: Dynamics of mitochondrial DNA evolution in mammals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. USA **86,** 6196–6200.

Kumar, S.; Tamura, K.; Nei, M., 1993: MEGA: molecular evolutionary genetics analysis, Version 1.01. University Park, PA: Inst. Mol. Evol. Genet. Penn. State University.

Lento, G. M.; Hickson, R. E.; Chambers, G. K.; Penny, D., 1995: Use of spectral analysis to test hypotheses on the origin of pinnipeds. Mol. Biol. Evol. **12,** 28–52.

Lockhart, P. J.; Steel, M. A.; Hendy, M. D.; Penny, D., 1994: Recovering evolutionary trees under a more general model of sequence evolution. Mol. Biol. Evol. **11,** 605–612.

Norman, J. E.; Ashley, M. V 2000: Phylogenetics of Perissodactyla and tests of the molecular clock. J. Mol. Evol. **50,** 11–21.

Palumbi, S., 1991: Simple Fool's Guide to PCR, Version 2.0. Town: Dept. Zool. Kewalo Marine Lab, Hawaii, pp. 32–33.

Pitra, C.; Kock, R. A.; Hofmann, R. R.; Lieckfeldt, D., 1998: Molecular phylogeny of the critically endangered Hunter's antelope (*Beatragus hunteri*). J. Zool. Syst. Evol. Research **36,** 179–184.

Prothero, D. R.; Manning, E.; Hanson, C. B., 1986: The phylogeny of the Rhinocerotoidea (Mammalia, Perissodactyla). Zool. J. Linn. Soc. **87,** 41–366.

Prothero, D. R.; Schoch, R. M., 1989: Origin and evolution of the Perissodactyla: summary and synthesis. In: Prothero, D. R.; Schoch, R. M. (eds), The Evolution of Perissodactyls. New York: Oxford University Press.

Radinsky, L. B., 1967: A review of the family Hyracodontidea (Perissodactyla). Bull. Am. Mus. Nat. Hist. **136,** 1–45.

Radinsky, L. B., 1969: The early evolution of the Perissodactyla. Evolution **23,** 308–328.

Sanger, F.; Nicklen, S.; Coulson, A. R., 1977: DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74,** 5436–5467.

Schoch, R. M., 1989: A review of the tapiroids. In: Prothero, D. R.; Schoch, R. M. (eds), The Evolution of Perissodactyls. New York: Oxford University Press.

Strimmer, K.; von Haeseler, A., 1996: Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. Mol. Biol. Evol. **13,** 964–969.

Thompson, J. D.; Higgins, D. G.; Gibson, T. J., 1994: CLUSTAL W:

improving the sensitivity of progressive multile sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res. **22,** 4673–4680.

Wägele, J.-W.; Rødding, F., 1998: A priori estimation of phylogenetic information conserved in aligned sequences. Mol. Phyl. Evol. **9,** 358–365.

Wood, H. E., 1937: Perissodactyl suborders. J. Mamm. **18,** 106–118.

Xu, X.; Arnason, U., 1994: The complete mitochondrial DNA sequence of the horse, *Equus caballus*: extensive heteroplasmy of the control region. Gene **148,** 357–362.

Xu, X.; Janke, A.; Arnason, U., 1996: The complete mitochondrial DNA sequence of the greater Indian rhinoceros, *Rhinoceros unicornis*, and the phylogenetic relationship among Carnivora, Perissodactyla, and Artiodactyla (+ Cetacea). Mol. Biol. Evol. **13,** 1167–1173.

Xu, X.; Gullberg, A.; Arnason, U., 1996: The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA comparisons among four closely related mammalian species-pairs. J. Mol. Evol. **43,** 438–446.

Xu, X.; Arnason, U., 1997: The complete mitochondrial DNA sequence of the white rhinoceros, *Ceratotherium simum*, and comparison with the mtDNA sequence of the Indian rhinoceros, *Rhinoceros unicornis*. Mol. Phylogenet. Evol. **7,** 189–194.

Zhang, D.-X.; Hewitt, G. M., 1996: Nuclear integrations: challenges for mitochondrial DNA markers. TREE **11,** 247–251.

Authors' address: C. Pitra (for correspondence), Jutta Veits, Institut für Zoo-und Wildtierforschung, FG Evolutionsgenetik. Alfred-Kowalke-Straße 17, 10315 Berlin, Germany. E-mail: Pitra@IZW-Berlin.de