

# Confirmação da relação entre peixes-bois, "hyraxes" e elefantes, por meio do estudo da proteína das lentes dos olhos(\*)

W. W. de Jong (\*\*)

A. Zweers (\*\*)

## Resumo

As relações filogenéticas entre os membros da super-ordem Paenungulata, os peixes-bois, elefantes e "hyraxes", foram investigadas, utilizando-se a seqüência de análises de aminoácidos da proteína  $\alpha$  A do cristalino da lente dos olhos destes animais. As análises confirmam a origem monofilética das ordens Hyracoidea, Sirenia e Proboscidea e colocam os edentatas como os mais antigos do ramo da linha principal dos euterianos, seguidos pelos Paenungulatas como ramo mais próximo. Os resultados sugerem que os peixes-bois podem estar mais próximos dos hyraxes que dos elefantes dada a presença dos 72-Leu nos hyraxes e peixes-bois, mas não nos elefantes. Evidências morfológicas indicariam uma relação mais próxima entre o elefante e peixe-boi, as quais, se verdadeira, exigirão uma retrossubstituição do 72-Leu  $\rightarrow$  Val na cadeia  $\alpha$  A dos elefantes ou 2 substituições paralelas na 72-Val  $\rightarrow$  Leu nos hyraxes e peixes-bois.

## UTILIZAÇÃO DA SEQUÊNCIA PROTÉICA EM ESTUDOS FILOGENÉTICOS

Bem como características morfológicas e anatômicas podem ser utilizadas para avaliar semelhanças entre espécies, e daí inferir relações filogenéticas, é possível comparar para esse propósito as propriedades estruturais de proteínas homólogas em diferentes espécies. Isso pode ser feito através de técnicas imunológicas, onde o grau de reatividade cruzada é uma medida da distância evolutiva entre espécies. Similarmente, um estudo da mobilidade eletroforética de proteínas pode fornecer informações sobre grau de parentesco entre espécies. Extremamente rica em informação, mas também difícil de realizar, é a comparação das seqüências de aminoácidos de proteínas homólogas de espécies cujas relações se

quer estudar. Isto não somente fornece uma medida concreta do grau de diferença entre duas proteínas, mas também estabelece as posições precisas e a natureza das substituições de aminoácidos. Isto possibilita, em princípio, a distinção na seqüência protéica de "resíduos" de aminoácidos primitivos e derivados de grande vantagem para a reconstrução segura das relações filogenéticas.

As técnicas de análise de seqüências de aminoácidos têm sido desenvolvidas e aperfeiçoadas pela química de proteínas, durante os últimos vinte anos. Pela relativa dificuldade, custo e consumo de tempo dessas técnicas, elas só têm sido aplicadas muito ocasionalmente a estudos filogenéticos específicos. Mais notáveis, no campo da filogenia de mamíferos, são os estudos sobre mioglobina (Romero-Herrera *et. al.*, 1978), hemoglobina (Goodman *et. al.*, 1979), citocromo c (Moore *et. al.*, 1976) e ribonuclease pancreática (Beintema *et. al.*, 1977). Esses estudos têm mostrado que comparações das seqüências de aminoácidos podem fornecer informações significativas sobre as relações filogenéticas entre taxa de mamíferos. Como foi mencionado por McKenna (1979), essas investigações, infelizmente, têm visado pouco à solução de casos de problemas filogenéticos, tais como as relações das pequenas e aberrantes ordens de mamíferos Edentata (preguiças, tamanduás, tatus), Pholidota (pangolins), Tubulidentata ("aardvarks", Proboscidea (elefantes), Hyracoidea ("hyraxes") e Sirenia (peixes-bois) entre si e com as ordens de mamíferos mais conhecidas.

(\*) — Em colaboração com o Projeto Peixe-Boi, do Departamento de Biologia de Mamíferos Aquáticos (INPA), Manaus.

(\*\*) — Department of Biochemistry, University of Nijmegen, 6500 HB Nijmegen, The Netherlands.

Estudos feitos em nosso laboratório mostraram no passado que comparações de seqüências da proteína  $\alpha$  cristalina das lentes dos olhos, ou cristalino, de 17 espécies, representando onze ordens de mamíferos, possibilitaram a construção de "árvores evolutivas" que concordam, em grande parte, com as opiniões geralmente aceitas sobre as relações entre essas ordens (De Jong *et. al.*, 1977). Ficou demonstrado, entre outras coisas, que elefantes e "hyraxes" são, mutuamente, os parentes mais próximos dentre os mamíferos estudados como de fato aceito pela maioria dos autores (Romer, 1966; Thenius, 1969, e outros) mas questionado por outros (McKenna, 1975). Notavelmente, em base à evidência da seqüência da  $\alpha$  cristalina, parece que essas duas ordens não são derivadas de uma origem comum com as ordens de ungulados, como geralmente aceito, mas representam antes um dos ramos mais antigos do tronco placentário principal.

Embora os Sirenia geralmente sejam considerados como muito proximamente relacionados aos Proboscidea e sejam, junto com os Hyracoidea, referidos como Paenungulata pareceu útil incluir o peixe-boi no estudo comparativo das  $\alpha$  cristalina de mamíferos. Isso não só possivelmente confirmaria (ou contradiria) as relações dos peixes-bois com os outros Paenungulata, como também aumentaria a "densidade" da árvore genealógica da  $\alpha$  cristalina melhorando, desse modo, a confiabilidade da informação filogenética deduzida em outros casos mais controvertidos.

#### $\alpha$ CRISTALINA: PROPRIEDADES E ANÁLISE DA SEQÜÊNCIA

A  $\alpha$  cristalina é uma proteína estrutural solúvel em água, que ocorre exclusivamente nas células do cristalino dos vertebrados. Compõe uma proporção variável das proteínas totais do cristalino, dependendo da idade e da espécie do animal. Em muitas espécies, ele atinge valores de 25 a 50% das proteínas totais (para uma revisão do assunto, vide Clayton, 1974).<sup>\*</sup> A  $\alpha$  cristalina forma grandes agregados, de peso molecular médio de 400.000 a 800.000, e é composta de dois tipos

de cadeias,  $\alpha$  A e  $\alpha$  B, que ocorrem numa proporção de aproximadamente 3:1 na maioria dos mamíferos placentários. As cadeias  $\alpha$  A e  $\alpha$  B do boi, que têm um comprimento de 173 a 175 "resíduos", respectivamente, mostram 55% de homologia em suas seqüências de aminoácidos, refletindo, assim, um artigo ancestral comum de seus genes.

A  $\alpha$  cristalina pode ser facilmente obtida em quantidades consideráveis na maioria das espécies de vertebrados e a seqüência de aminoácidos da cadeia  $\alpha$  A é relativamente simples de estabelecer-se. De estudos anteriores, (De Jong *et. al.*, 1977) poderia estimar-se que a cadeia  $\alpha$  A desenvolve uma taxa de 2,9 substituições de aminoácidos por 100 "resíduos" em 100 milhões de anos. Isso a coloca entre as proteínas de mais lenta evolução e a torna adequada para o estudo das relações filogenéticas entre taxa de ordens superiores.

Os procedimentos gerais envolvidos na determinação da seqüência das cadeias da  $\alpha$  cristalina A são descritos em detalhe em De Jong e Terwindt (1976) e serão dados aqui resumidamente, como aplicados à elucidação da seqüência da cadeia da  $\alpha$  cristalina A do peixe-boi. Seis cristalinos de peixes-bois (**Trichechus inunguis**) jovens do Brasil foram obtidos do Dr. R.C. Best, Instituto Nacional de Pesquisas da Amazônia, Manaus. Os cristalinos foram enviados por via aérea, à temperatura ambiente, numa solução saturada de cloridrato de guanidina. Sob essas condições, os cristalinos são completamente solubilizados, mas as proteínas ficam protegidas de ataque microbiano ou enzimático e permanecem (ao contrário de quando conservada em álcool ou formol) adequadas para posterior análise estrutural. A cadeia de  $\alpha$  cristalina A podia ser recuperada dessa solução e separada de todas as outras proteínas do cristalino por cromatografia de coluna de troca iônica, depois de removido o cloridrato de guanidina por diálise. A cadeia  $\alpha$  A isolada foi digerida com tripsina que especificamente corta a cadeia atrás dos aminoácidos básicos lisina e arginina. A resultante mistura de cerca de 20 peptídeos pequenos foi fracionada por cromatogra-

fia de filtração em gel e mapeamento peptídico (eletroforese de alta voltagem em papel, seguida por cromatografia em papel na segunda dimensão). Alguns peptídeos tripsínicos ("tryptic peptides") maiores foram digeridos posteriormente com termolisina. Foram feitas análises de aminoácidos do peptídeos purificados e as composições comparadas com os peptídeos correspondentes das cadeias  $\alpha$  A de bois e elefantes, que já haviam sido seqüenciadas. Quando a carga e composição dos pequenos peptídeos homólogos eram os mesmos, admitia-se que suas seqüências de

aminoácidos também eram idênticas.

Desde que esse procedimento de análise de seqüências de aminoácidos se fia, num alto grau, na comparação das composições de aminoácidos de peptídeos pequenos, há obviamente o risco de omitir substituições duplas de tipo oposto, ocorrendo no mesmo peptídeo sem mudar sua composição. Tem sido estabelecido, contudo, que esse risco de omitir tais substituições recíprocas é desprezível, contanto que os peptídeos analisados sejam suficientemente pequenos (Van Druten *et. al.*, 1978).

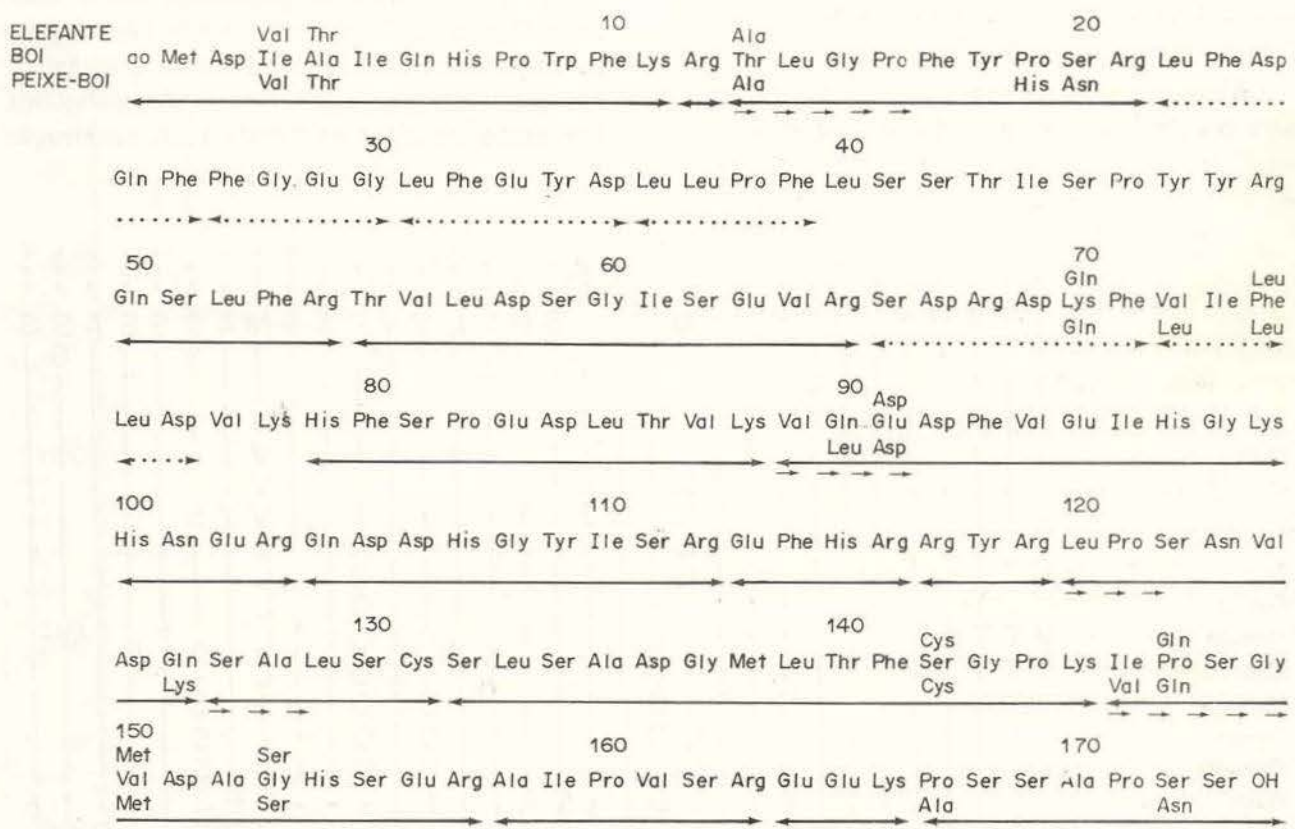


Fig. 1 — Resumo da análise da seqüência da cadeia da  $\alpha$  cristalina A do peixe-boi. A linha central representa a seqüência anteriormente bem estabelecida, da cadeia  $\alpha$  A bovina. Os dez radicais que foram considerados diferentes da cadeia  $\alpha$  A do elefante (De Jong *et al.*, 1977) são mostrados acima da seqüência bovina e as 18 diferenças agora encontradas na cadeia  $\alpha$  A do peixe-boi são mostradas abaixo dessa seqüência. São indicados os peptídeos tripsínicos ("tryptic") ( $\leftarrow\rightarrow$ ) e termilíticos ( $\leftarrow\cdots\rightarrow$ ), os quais podem ser isolados da cadeia  $\alpha$  A de peixe-boi e que foram usados para deduzir a seqüência proposta. Nenhum peptídeo correspondente às posições 40-49 e 77-78 pode ser isolado da cadeia  $\alpha$  A do peixe-boi; essas posições são, contudo, idênticas à seqüência bovina, em todas as espécies de mamíferos estudadas. As posições de um número de "resíduos" foram determinadas pela degradação química seqüencial dos peptídeos, e assim chamada degradação de Edman ( $\rightarrow\rightarrow$ ). A maioria dos outros "resíduos" estão ordenados, assumindo-se a homologia seqüencial dos peptídeos bovinos e do elefante. A substituição Ser  $\rightarrow$  Asn ocorreu na posição 172 ou 173.

SEMELHANÇA DA CADEIA  $\alpha$  A DO PEIXE-BOI  
COM AS DO "HYRAX" E DO ELEFANTE

Verificou-se que a seqüência  $\alpha$  A do peixe-boi compartilha de todas as dez substituições que distinguem a cadeia  $\alpha$  A do elefante da do foi (Figura 1). A cadeia  $\alpha$  A do peixe-boi diferiu em 8 posições da cadeia  $\alpha$  A do elefante; uma delas, 72-Leu, é encontrada somente na cadeia  $\alpha$  A do "hyrax". As posições das 5 substituições que aparecem unicamente na cadeia  $\alpha$  A do peixe-boi foram deduzidas da especificidade da separação enzimática (126-Lys, 167-Ala), da diferença em composição com o peptídeo (19-His, 20-Asn) homólogo do elefante, ou não puderam ser estabelecidas com certeza (172 ou 173-Asn).

A comparação da seqüência  $\alpha$  A proposta para o peixe-boi com o conjunto de dados das

cadeias  $\alpha$  A de outros mamíferos estudados, como mostra a Tabela 1, facilmente revela que, de fato, as cadeias  $\alpha$  A do peixe-boi, "hyrax" e elefante compartilharam de um certo número de substituições únicas. Parece que as substituições 70 Lys  $\rightarrow$  Gln, 72 Val  $\rightarrow$  Leu, 74 Phe  $\rightarrow$  Leu e 142 Ser  $\rightarrow$  Cys são características para as seqüências  $\alpha$  A de Paenungulata, embora as duas últimas seqüências também ocorram como substituições independentes e paralelas no tamanduá e no homem, respectivamente.

A seqüência  $\alpha$  A do peixe-boi pode ser conectada à árvore de cadeias  $\alpha$  A de 17 mamíferos, anteriormente publicada, como mostra a Figura 2. Essa árvore ainda não contém as seqüências mais recentemente analisadas, das preguiças, do tamanduá e do pangolim, que estão incluídas na Tabela 1. A construção

	3	4	13	16	17	56	58	61	70	72	74	90	91	93	101	112	113	117	122	124	126	134	137	144	145	150	152	153	155	156	158	167	168	169	
Baleia	I	A	A	P	F	V	D	I	K	V	F	Q	E	F	N	S	N	S	L	S	V	P	S	G	M	A	G	S	E	A	S	S	S		
Cavalo	I	A	A	P	F	V	D	I	K	V	F	Q	E	F	N	S	N	S	L	S	V	P	S	G	M	A	G	S	E	A	S	S	G		
Rinoceronte	I	A	A	P	F	V	D	I	K	V	F	Q	E	F	N	S	N	S	L	S	V	P	S	G	M	A	G	S	E	A	S	S	G		
Porco	I	A	A	P	F	V	D	I	K	V	F	Q	E	F	N	S	N	S	L	S	V	P	S	G	M	A	G	S	E	A	S	S	G		
Boi	I	A	A	P	F	V	D	I	K	V	F	Q	E	F	N	S	N	S	L	S	V	P	S	G	M	A	G	S	E	A	S	S	G		
Cão, Gato	I	A	A	P	F	V	D	I	K	V	F	Q	E	F	N	S	N	S	L	S	V	P	S	G	M	A	G	S	E	A	S	S	G		
Pangolim	I	A	A	P	F	V	D	I	K	V	F	Q	E	F	N	S	N	S	L	S	V	P	S	G	M	A	G	S	E	A	S	S	G		
"Hedgehog"	V	T	T								Y	L	L	S	S	P	S	A																	
Rato, "Tupaia"	V	T	T								Y	L	L	S	S	P	S	A																	
Coelho	V	T	T								Y	L	L	S	S	P	S	A																	
Homem	V	T	T											D									C	I	Q	T							T		T
Macaco Rhesus	V	T	T											D									C	I	Q	T							T		T
Elefante	V	T						Q		L		D											C	I	Q				S						
"Hyrax"	V	T						Q	L	L	L	D											C	I	Q				S						
Peixe-Boi	V	T						Q	L	L	L	D											C	I	Q				S						
Preguiça Real	V	T										L	D				T	A					I	-	-	-	V	P	S				T		
Preguiça comum	V	T				A						L	G				T	A					I	-	-	-	V	P	S				P		
Tamanduá	V	T				A		R		L	L	D				T	A						L	-	-	-	V	P	S	G			T		
Canguru	I	T		S	L	E		R		I	L	D		S								I	I	H	I	D		I	S		D	S	T	L	
Gambá	I	T		S	L	E		R		Y	L	D	Y	S								I	I	H	T	N		S		D	S	T	L		

TABELA 1 — Resumo das posições filogeneticamente informativas das seqüências  $\alpha$  A de mamíferos, isto é, posições cujas substituições ocorrem em, pelo menos, duas espécies investigadas. Os nomes completos das espécies podem ser encontrados em De Jong et al., 1977 e 1981. As linhas verticais indicam onde os radicais são idênticos à seqüência mais alta. O ramo Paenungulata é determinado pelas substituições nas posições indicadas por setas ( $\uparrow$ ). A notação de uma só letra, para aminoácidos, foi utilizada A=Ala; C=Cys; D=Asp; E=Glu; F=Phe; G=Gly; H=His; I=Ile; K=Lys; L=Leu; M=Met; N=Asn; P=Pro; Q=Gln; R=Arg; S=Ser; T=Thr; V=Val; Y=Tyr; (-) indica anulação.

das árvores filogenéticas desse conjunto de seqüências  $\alpha$  A por computador, será apresentada em outra publicação (De Jong *et. al.*, 1981). Pode mencionar-se também que essa análise confirma a origem monofilética dos ordens Hyracoidea, Sirenia e Proboscidea e situa os Edentata como o mais velho ramo da linha principal dos Eutheria, seguidos pelos Paenungulata como o ramo seguinte.

A árvore apresentada, na Figura 2, representa os elefantes como sendo o primeiro ra-

mo dos Paenungulata, sugerindo uma relação mais próxima entre o "hyrax" e o peixe-boi. Isso se deve à presença de 72-Leu na cadeia  $\alpha$  A do "hyrax" e do peixe-boi mas não na do elefante. Há algumas indicações morfológicas de que, na realidade, o peixe-boi está mais proximamente relacionado aos elefantes do que aos "hyraxes" (McKenna, 1975). Se isso for verdade, deve-se admitir uma substituição retrógrada de 72 Leu  $\rightarrow$  Val na cadeia  $\alpha$  A do elefante, ou duas substituições paralelas 72

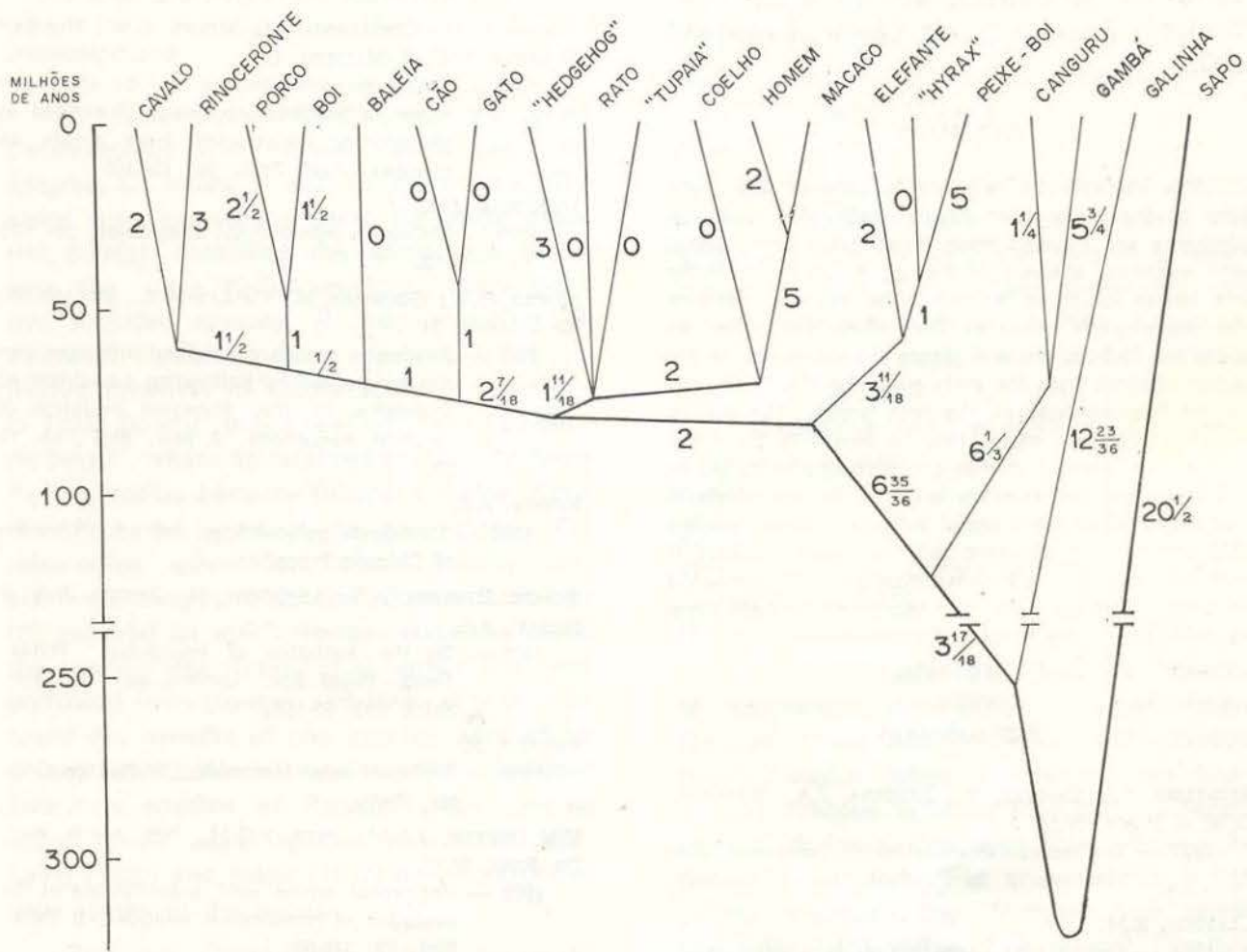


Fig. 2 — Árvore filogenética das cadeias do  $\alpha$  cristalina A. A seqüência  $\alpha$  A do peixe-boi foi adicionada em base dos dados fornecidos na Tabela I, à árvore anteriormente publicada (De Jong *et al.*, 1977). Essa árvore foi construída, através do uso de computador a partir das diferenças entre seqüências de cadeias  $\alpha$  A de mamíferos, minimizando o número de substituições necessárias à sua evolução. Todos os outros padrões da ramificação entre as cadeias  $\alpha$  A dos mamíferos poderão requerer mais substituições do que as adotadas nessa árvore. Os comprimentos dos ramos estão expressos nas substituições; os valores fracionais aparecem onde há dúvida, como para o radical do ancestral no nó. Os tempos de divergência (em milhões de anos) são estimados com base em dados paleontológicos.

Val → Leu no "hyrax" e no peixe-boi, para conseqüentemente, mudar o padrão de ramificação do Paenungulata mostrada na Figura 2.

#### AGRADECIMENTOS

Somos muito gratos ao Dr. R. C. Best (INPA), Manaus, por fornecer as lentes de peixes-bois. A Srta. Marlies Versteeg amavelmente realizou as análises de vários aminoácidos e Megumi Yamakoshi (INPA), Manaus, traduziu o texto para o português. Este trabalho foi patrocinado, em parte, pela Netherlands Foundation for Chemical Research (SON).

#### SUMMARY

The phylogenetic relationship between the members of the super-order Paenungulata, the manatees, elephants and hyraces were investigated using amino acid sequence analysis of the  $\alpha$  A crystallin of the eye lenses of these animals. The analysis confirms the monophyletic origin of the orders Hyracoidea, Sirenia and Proboscidea and places the edentates as the oldest offshoot from the main eutherian stem, followed by the Paenungulates as the next branch. The results suggest that the manatee may be closer to the hyrax than to the elephant due to the presence of 72-Leu in the hyrax and the manatee but not in the elephant. Morphological evidence would indicate a closer relationship between the elephant and the manatee which, if true, would require a back-substitution of 72 Leu → Val in the elephant  $\alpha$  A chain or two parallel substitutions 72 Val → Leu in the hyrax and manatee.

#### BIBLIOGRAFIA

- BEINTEMA, J.J.; CAASTRA, W.; LENSTRA, J.A.; WELLING, G.W. & FITCH, W.M.  
1977 — The molecular evolution of pancreatic ribonuclease. *J. Mol. Evol.*, 10: 49-71.
- CLAYTON, R.M.  
1974 — Comparative aspects of lens proteins, in *The Eye*. Vol. 5 (H. Davson and L.T. Graham, eds.), Academic Press, New York, p. 399.
- DE JONG, W.W.; GLEAVES, J.T. & BOULTER, D.  
1977 — Evolutionary changes of  $\alpha$  crystallin and the phylogeny of mammalian orders. *J. Mol. Evol.*, 10: 123-135.
- DE JONG, W.W. & TERWINDT, E.C.  
1976 — The amino-acid sequences of the  $\alpha$  — crystallin A chains of red kangaroo and Virginia opossum. *Eur. J. Biochem.*, 67: 503-510.
- DE JONG, W.W.; ZWEERS, A.; JOYSEY, K.A.; GLEAVES, J.T. & BOULTER, D.  
1980 — Protein sequence analysis applied to edentate phylogeny, in *The Evolution and Ecology of Sloths, Anteaters and Armadillos* (G. G. Montgomery, ed.), Smithsonian Institution Press, Washington, D.C. (in press).
- GOODMAN, M.; CZELLUSNIAK, J.; MOORE, G.W.; ROMERO-HERRERA, W.E. & MATSUDA, G.  
1979 — Fitting the gene lineage into its species lineage, a parsimony strategy illustrated by cladograms constructed from globin sequences. *Syst. Zool.*, 28: 132-163.
- MCKENNA, M.C.  
1979 — Molecular mammalogy. *Syst. Zool.*, 28: 109-113.
- MOORE, G.W.; GOODMAN, M.; CALLAHAN, C.; HOLMQUIST, K. & MOISE, H.  
1976 — Stochastic versus augmented maximum parsimony method for estimating superimposed mutations in the divergent evolution of protein sequences. *J. Mol. Biol.*, 15: 15-37.
- ROMER, A.S.  
1966 — *Vertebrate paleontology*, 3rd ed., University of Chicago Press.
- ROMERO-HERRERA, A.E.; LEHMANN, H.; JOYSEY, K.A. & FRIDAY, A.E.  
1978 — On the evolution of myoglobin. *Philos. Trans. Royal Soc. London*, ser B (Biol. Sci.), 283: 63-163.
- THENIUS, E.  
1969 — *Phylogeni eder Mammalia*. Walter de Gruyter, Berlin.
- VAN DRUTEN, J.A.M.; PEER, P.G.M.; BOS, A.B.H. & DE JONG, W.W.  
1978 — Reciprocal amino acid substitutions in the evolution of homologous peptides. *J. Theor. Biol.*, 73: 549-561.

(Aceito para publicação em 19/06/80)