FEBS 16805 FEBS Letters 382 (1996) 319–322

Pseudogenes in ribonuclease evolution: a source of new biomacromolecular function?

Nathalie Trabesinger-Ruef[a], Thomas Jermann[a], Todd Zankel[a], Barbara Durrant[c], Gerhard Frank[b], Steven A. Benner[a,*]

[a] *Department of Chemistry, ETH, CH-8092 Zürich, Switzerland*
[b] *Institute for Molecular Biology, ETH, CH-8093 Zürich, Switzerland*
[c] *Center for the Reproduction of Endangered Species, Zoological Society of San Diego, San Diego, CA 92112-0551 CA, USA*

Received 27 November 1995; revised version received 13 February 1996

Abstract Bovine seminal ribonuclease (RNase) diverged from pancreatic RNase after a gene duplication ca. 35 million years ago. Members of the seminal RNase gene family evidently remained as unexpressed pseudogene for much of its evolutionary history. Between 5 and 10 million years ago, however, after the divergence of kudu but before the divergence of ox, evidence suggests that the pseudogene was repaired and expressed. Intriguingly, detailed analysis of the sequences suggests that the repair may have involved gene conversion, transfer of information from the pancreatic gene to the RNase pseudogene. Further, the ratio of non-silent to silent substitutions suggests that the pancreatic RNases are divergently evolving under functional constraints, the seminal RNase pseudogenes are diverging under no functional constraints, while the genes expressed in the seminal plasma are evolving extremely rapidly in their amino acid sequences, as if to fulfil a new physiological role.

Key words: Ribonuclease; Evolution; Gene conversion; Artiodactyls; Seminal plasma

1. Introduction

New biomolecular function is believed to arise, at least in recent times, largely through recruitment of established proteins to play new roles following gene duplication [1,2]. Under one model, one copy of a gene continues to divergently evolve under constraints dictated by the ancestral function. The duplicate, meanwhile, is unencumbered by a functional role, and is free to search protein "structure space". It may eventually come to encode new behaviors required for a new physiological function, and thereby confer selective advantage.

This model contains a well recognized paradox. Because duplicate genes are not under selective pressure, they should also accumulate mutations that render them incapable of encoding a protein useful for any function. Most duplicates therefore should become pseudogenes, inexpressible genetic information ("junk DNA" [3]) in just a few million years [4,5]. This limits the evolutionary value of a functionally unconstrained gene duplicate as a tool for exploring protein "structure space" in the search of new behaviors that might confer selectable physiological function.

Ribonucleases (RNases) offer an interesting system to study how new function arises in proteins [6]. RNase A is a well known member of this family expressed in the pancreas of ox, where it serves a role in the intestine digesting RNA arising from bacteria fermenting in the first stomach of the ox [7]. Ox also contains a homologous RNase, 23 amino acids different, expressed in the seminal plasma, where it constitutes some 2% of total protein [8]. Seminal RNase has evolved to become a dimer with composite active sites. It binds tightly to anionic glycolipids [9–12], including seminolipid, a fusogenic sulfated galactolipid found in bovine spermatozoa [13]. The ancestral RNase does not bind seminolipid. Further, seminal RNase has immunosuppressive and cytostatic activities that are not displayed by the ancestral RNase [14,15].

Laboratory reconstructions of ancient RNases [16] have suggested that each of these traits was not present in the most recent common ancestor of seminal and pancreatic RNase, but rather arose in the seminal lineage after the divergence of these two protein families. To learn more about how this remarkable example of evolutionary recruitment occurred, we have analyzed RNase genes from peccary (*Tayassu pecari*), Eld's deer (*Cervus eldi*), domestic sheep (*Ovis aries*), oryx (*Oryx leucoryx*), saiga (*Saiga tatarica*), yellow-backed duiker (*Cephalophus sylvicultor*), lesser kudu (*Tragelaphus imberbis*) and Cape buffalo (*Syncerus caffer caffer*), diverging approximately in that order within the mammal order *Artiodactyla* [17,18]. These complement the genes for ox pancreatic RNase [19], seminal RNase from ox (*Bos taurus*) [20], and giraffe (*Giraffa camelopardalis*) [21] and hog deer (Beintema, personal communication) (Figs. 1 and 2).

2. Materials and methods

Spermatozoa were isolated by centrifugation from the respective seminal plasmas, obtained from the collection of the Center for Reproduction of Endangered Species and incubated overnight in a buffer containing SDS and proteinase K [22]. The derived genomic DNA was amplified in a 50 μl reaction volume containing: 2 mM $MgCl_2$, 50 mM KCl, 10 mM Tris/HCl (pH 9.0), 0.1% Triton X-100, 0.2 mM dNTP's, 2 U Taq polymerase (Promega), 50 pmol primers hybridizing to codons 5–14 in the 5'-region and to codons 119–124 at the 3'-end of BS-RNase, and 200 ng genomic DNA. PCR reaction conditions: 1 min 95°C, 2 min 52°C, 50 s 72°C, 3 cycles and 8 min 72°C [23]. The PCR fragments of saiga and kudu were cloned in a pUC19 plasmid using T4 DNA Ligase (NEB) in a 1× ligation buffer: 50 mM Tris/HCl (pH 7.8), 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 μg/ml BSA at 14°C overnight [19]. The PCR products were analysed either by direct sequencing (USB-kit) or by start and reverse primers (Synthesizer ABI, 380A) [24].

Seminal plasma was also tested for RNase activity using a zymogram spot assay [25]. Sample (1 μl) was placed on a 1% agarose gel (10 mM Tris-HCl, pH 7.5) containing poly(C) (0.3 mg/ml). The gel was incubated at 37°C for 30 min, then stained with Tris-HCl buffer (10 mM, pH 7.5) containing toluidine blue (2 mg/ml), and then destained with water. The concentration was calculated assuming a specific activity identical to RNase A in these assays. Proteins in the seminal plasma were also resolved by gel electrophoresis and identified by Western blotting [19].

*Corresponding author. Fax: (41) (1) 632 1170.
E-mail: benner@chem.ethz.ch

SSDI S0014-5793(96)00191-3

3. Results and discussion

Seminal RNase genes are distinguished from their pancreatic cousins by several "marker" substitutions introduced early after the gene duplication, including Pro 19, Cys 32, and Lys 62. By this standard, as well as by parsimony analysis, the genes from saiga, sheep, duiker, kudu, and Cape buffalo were all assigned to the seminal RNase family. No evidence for a seminal-like gene could be found in peccary. Thus, these data are consistent with a parsimony analysis of previously published genes that places the gene duplication separating pancreatic and seminal RNases ca. 35 million years before present [6], and the divergence of giraffe preceding the divergence of sheep, saiga, duiker, kudu, Cape buffalo and ox, in this order, consistent with mitochondrial sequence data [18].

Sequence analysis shows that the seminal RNase genes from both kudu and Cape buffalo almost certainly could not serve physiological functions by encoding a folded stable protein. A single base deletion disrupts codon 114 in kudu, creating a frame shift. A deletion introducing a frame shift and involving codons 54–57 is found in the Cape buffalo gene. This implies that these are pseudogenes. Breukelmann et al. [21] noted that the seminal RNase gene in giraffe might also be a pseudogene, based on the rate of sequence divergence. Lesions have now evidently been found in these two genes as well (Beintema, personal communication).

To show that these seminal genes were indeed not expressed in semen, plasmas from 15 artiodactyls were examined (ox, forest buffalo (*Syncerus caffer nanus*), Cape buffalo, kudu, sitatunga (*Tragelaphus spekei*), nyala (*Tragelaphus angasi*), eland (*Tragelaphus oryx*), Maxwell's duiker (*Cephalophus monticola maxwelli*), yellow-backed duiker, suni (*Neotragus*

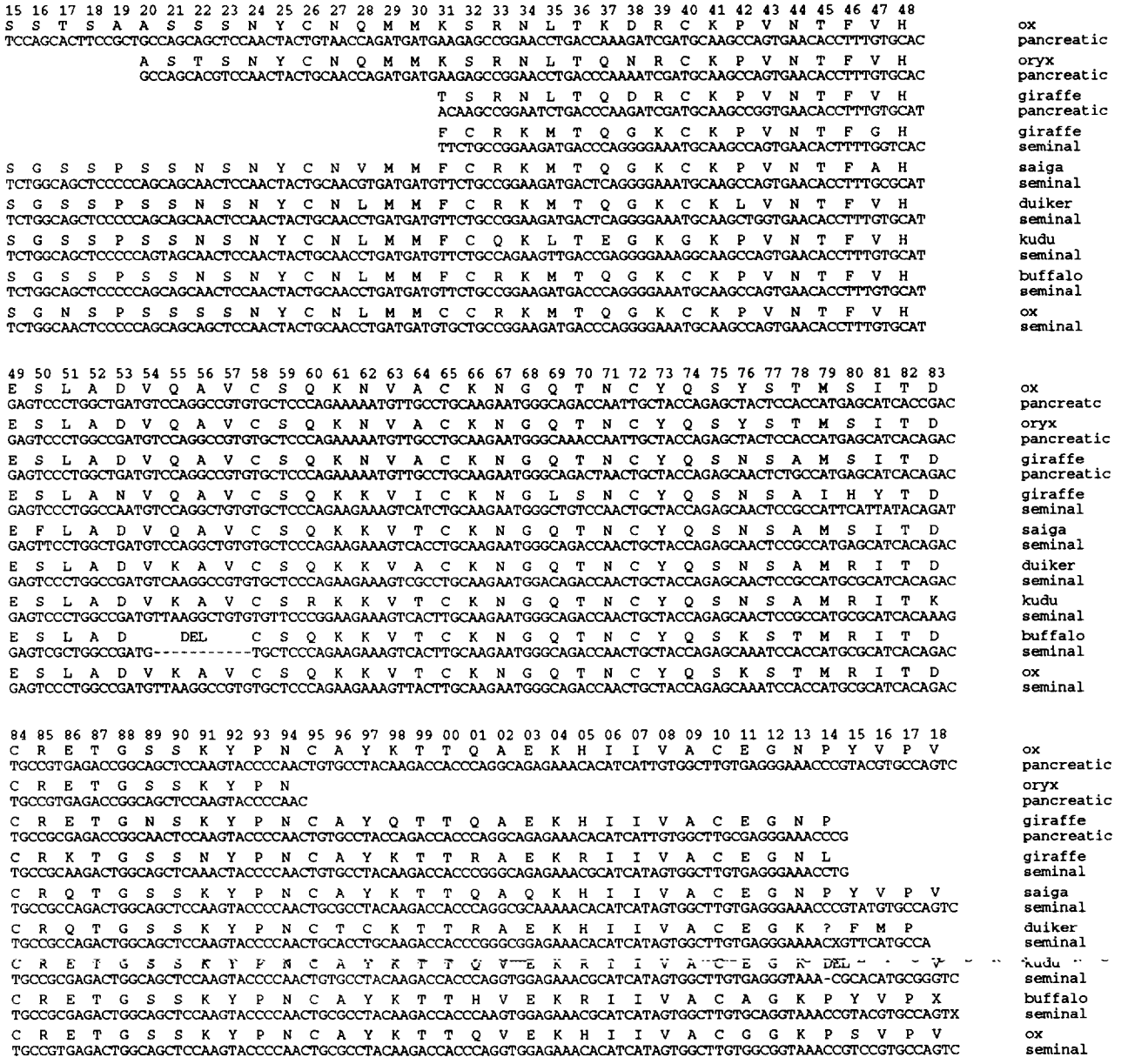

Fig. 1. Sequences of the genes and the encoded polypeptide chain (using the one letter code) of members of the pancreatic and seminal RNase families. Published RNase sequences are from giraffe [21], ox pancreas [19] and ox seminal plasma [20].

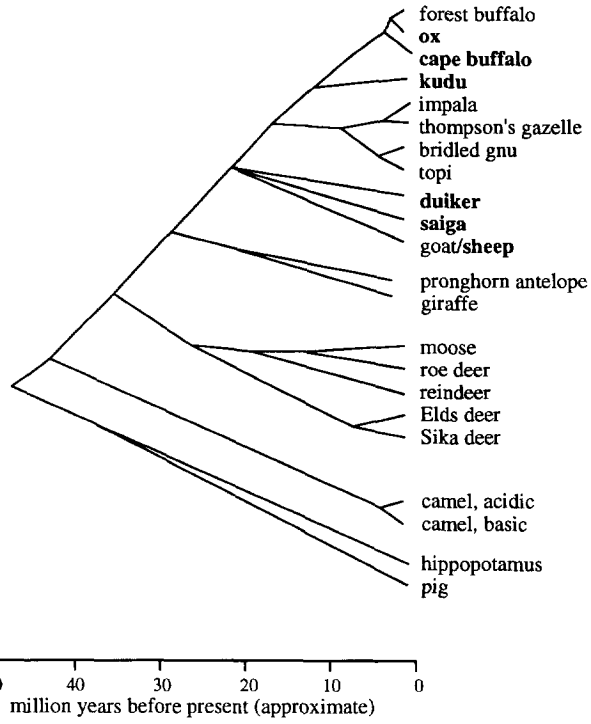

Fig. 2. A phylogenetic tree showing the relationship between artiodactyl (even-toed ungulate) organisms discussed here. The organisms contributing sequences discussed in the text are shown in bold.

moschatus), sable antelope (*Hippotragus niger*), impala (*Aepyceros melampus*), saiga (*Saiga tatarica*), sheep (*Ovis aries*), and Eld's deer). Catalytically active RNase was not detected in the seminal plasma in significant amounts in any artiodactyl genus diverging before the Cape buffalo, except in *Ovis*. Independent mutagenesis experiments showed that the proteins encoded by these genes, all carrying a Cys at position 32, should form dimers [9–12]. By Western blotting, however, only small amounts of a monomeric, presumably pancreatic RNase, were detected in these seminal plasmas. In contrast, the seminal plasmas of sheep, forest buffalo, cape buffalo and ox all contained substantial amounts of Western blot-active RNase. The seminal plasma of ox was the only sample found to contain significant amounts of a band on a denaturing, non-reducing gel corresponding to the seminal RNase dimer.

This result was unexpected on paleontological grounds, because the *Ovis* genus diverged from the lineage leading to modern ox long before other genuses lacking seminal RNase diverged. To learn whether RNase in the *Ovis* seminal plasma is derived from a seminal RNase gene, the RNase from goat seminal plasma was isolated, purified, and sequenced by tryptic cleavage and Edman degradation. Both Edman degradation (covering 80% of the sequence) and MALDI mass spectroscopy showed that the sequence of the RNase isolated from goat seminal plasma is identical to the sequence of its pancreatic RNase (data not shown) [6,9]. This shows that the RNase in *Ovis* seminal plasma is not expressed from a seminal RNase gene, but rather from the *Ovis* pancreatic gene. To confirm this conclusion, a fragment of the seminal RNase gene from sheep was sequenced, and shown to be different in structure from the pancreatic gene (data not shown). The fact that the Western-blot active bands observed in Cape and forest buffalo had mobilities on a denaturing gel consistent with a glycosylated pancreatic monomeric RNase A, but not a dimeric seminal RNase as found in ox, suggests that these also might be derived from the pancreatic RNase A lineage.

These results might be consistent a model that assumes that seminal RNase gained a physiological function immediately after duplication, that this function was retained throughout the divergent evolution, is retained in modern ox, but was lost in all other species, including kudu and Cape buffalo. This would require, however, that this function was lost independently multiple times in different lineages.

More likely, however, is a model where the duplicate RNase gene initially served no function, and therefore suffered damage, such as that reflected in the modern kudu and Cape buffalo. The gene was then resurrected after the divergence of Cape buffalo in the lineage leading to modern ox. This resurrection must have been very recent. Clades containing the saiga, duiker and sheep are known in the early Miocene (23.8–16.4 million years before present), while clades containing the kudu and cape buffalo are known in the late Miocene (11.2–5.3 my bp), implying that the pseudogene acquired new function only within the past few million years, possibly more recently. It is intriguing to ask whether the domestication of the ox is related to the emergence of seminal RNase as a functioning protein.

The question then arises as to whether the seminal RNase gene indeed has a function in modern ox. An observation relevant to this question concerns the ratio of non-silent and silent substitutions in these gene families. In unexpressed seminal RNase sequences, the ratio of non-silent to silent substitutions averages 2:1. This is close to that expected for random substitution in a gene serving no selected function, and is consistent with the model that these seminal RNases are pseudogenes. In contrast, the average ratio is less than 1:1 with pancreatic RNases, consistent with the proposal that pancreatic RNases are functioning genes where amino acid replacements are constrained by selective pressures. Most remarkable, however, is the ratio of non-silent to silent substitutions observed when comparing the expressed ox seminal RNase with its nearest unexpressed homologs, from Cape buffalo and kudu. This ratio is ca. 4:1. This is expected only for a pseudogene emerging after searching protein "structure space" to perform a new function, with amino acid substitutions rapidly introduced to provide new selected properties. Interestingly, the introduction of Cys 31 is evidently associated with the resurrection of the seminal RNase gene.

How was this pseudogene resurrected? It is difficult at this point to say. It is interesting to note that in the region of the kudu deletion, the sequence of the expressed seminal RNase gene in ox is quite similar to the sequence of the ox pancreatic gene, more than it is to the kudu seminal RNase pseudogene, and that this similarity extends some 70 base pairs into the 3′-untranslated region (with 62 of the 70 nucleobases, 89%, identical). We may speculate that information from the pancreatic gene may have been used to repair a damaged ancestral seminal RNase gene, perhaps by a gene conversion event [26–29]. This may be the first example of gene conversion being used to create new physiological function in paleogeological evolution, and it will be interesting to test this hypothesis with more sequence data.

Acknowledgements: We are indebted to Dr. Alan Gentry, of the Brit-

ish Museum of Natural History, and Dr. Malcolm C. McKenna, Frick Curator in the Department of Vertebrate Palaeontology at the American Museum of Natural History, for helpful comments concerning artiodactyl evolution, and Prof. Jaap Beintema, for fascinating discussions and his many contributions to this field.

References

[1] Ohno, S. (1970) in: Evolution by Gene Duplication, Springer, NY.
[2] Benner, S.A. and Ellington, A.D. (1990) Bioorganic Chemistry Frontiers 1, 1–70.
[3] Li, W.H., Gojobori, T. and Nei, M. (1981) Nature 292, 237–239.
[4] Jukes, T.H. and Kimura, M. (1984) J. Mol. Evol. 21, 90–92.
[5] Marshall, C.R., Raff, E.C. and Raff, R.A. (1994) Proc. Natl. Acad. Sci. USA 91, 12283–12287.
[6] Beintema, J.J., Schuller, C., Irie, M. and Carsana, A. (1988) Prog. Biophys. 51, 165–192.
[7] Barnard, E.A. (1969) Nature 221, 340–344.
[8] D'Alessio, G. (1962) Biochem. J. 7 pp.
[9] Jermann, T.M. (1995) ETH Dissertation No. 11059.
[10] Opitz, J.G. (1995) ETH Dissertation No. 10952.
[11] Trauwein-Fritz, K. (1991) ETH Dissertation No. 9613.
[12] Raillard-Yoon, S.A. (1993) ETH Dissertation No. 10022.
[13] Vos, J.P., Lopes-Cardozo, M. and Gadella, B.M. (1994) Biochem. Biophys. Acta Lipid 1211, 125–149.
[14] Soucek, J., Chudomel, V., Potmesilova, I. and Novak, J.T. (1986) Nat. Immun. Cell Growth Regul. 5, 250–258.
[15] Benner, S.A. (1989) Trends Biochem. Sci. 14, 396–397.
[16] Jermann, T.M., Opitz, J.G., Stackhouse, J. and Benner, S.A. (1995) Nature 374, 57–59.
[17] Carroll, R.L. (1988) in: Vertebrate Paleontology and Evolution. W.H. Freeman, New York.
[18] Allard, M.W., Miyamoto, M.M., Jarecki, L., Kraus, F. and Tennant, M.R. (1992) Proc. Natl. Acad. Sci. USA 89, 3972–3976.
[19] Carsana, A., Confalone, E., Palmieri, M., Libonati, M. and Furia,A. (1988) Nucl. Acids Res. 16, 5491–5502.
[20] Preuss, K.-D., Wagner, S., Freudenstein, J. and Scheit, K.-H. (1992) Nucl. Acids Res. 18, 1057.
[21] Breukelman, H.J., Beintema, J.J., Confalone, E., Costanzo, C., Sasso, M.P., Carsana, A., Palmieri, M. and Furia, A. (1993) J. Mol. Evol. 37, 29–35.
[22] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) in: Molecular Cloning. A Laboratory Manual. 2nd Ed. Cold Spring Harbor Laboratory. Cold Spring Harbor, NY.
[23] Mullis, K.B. and Falcoona, F.A. (1987) Methods Enzymol. 155, 335–350.
[24] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[25] delCardayré, S.B., Ribo, M., Yokel, E.M., Quirk, D.J., Rutter, W.J. and Raines, R.T. (1995), Prot. Eng. 8, 261–273.
[26] Wysocki, L. and Gefter, M.L. (1989) Annu. Rev. Biochem. 58, 509–531.
[27] Eikenboom, J.C.J., Vink, T., Briet, E., Sixma, J.J. and Reitsma, P.H. (1994) Proc. Natl. Acad. Sci. USA 91, 2221–2224.
[28] Klein, J. (1984) Transplantation 38, 327–329.
[29] Hughes, A.L. (1991) Immunogenetics 33, 367–373.