

# Molecular cloning of exons II and III of the $\alpha$ -globin major gene from *Odontophrynus americanus* 2n and 4n (Amphibia, Anura)

Maria Dolores P. Acedo<sup>1</sup>, Gláucia Paranhos-Baccalà<sup>2</sup>,  
Claudio D. Denoya<sup>3</sup> and Itamar R.G. Ruiz<sup>4</sup>

## ABSTRACT

The  $\alpha$ -globin major genes from diploid and tetraploid *Odontophrynus americanus* were studied using PCR-based technology. The cloned and sequenced amplified fragments were shown to contain most of the exon II sequences as well as the whole exon III sequence of the  $\alpha$ -globin gene. Unexpectedly, intron 2 was entirely absent in the amplified fragments of both 2n and 4n origin. High conservation was observed among the obtained sequences when compared to corresponding sequences from human and *Xenopus laevis* origin. The possibility that these sequences might be pseudogenes is raised.

## INTRODUCTION

The sibling species *Odontophrynus americanus* 2n and 4n exist as endemic populations in South America. The 4n species probably arose through a recent polyploidization event (Beçak and Beçak, 1974). The 2n/4n system has been used as an excellent model for biochemical (Beçak and Goissis, 1971), cytogenetical (Ruiz *et al.*, 1981) and, especially, molecular evolution studies (Cortadas and Ruiz, 1988; Ruiz and Brison, 1989), and constitutes an important tool for the study of molecular divergence in closely related species. Studies performed on ribosomal cistrons showed a great polymorphism of the rDNA repeats among 2n and 4n specimens. The 5.8S and part of the 18S/28S rDNA sequences are highly conserved among 2n/4n and *X. laevis*

(Amphibia, Anura), whereas the internal transcribed spacers 1 and 2 (ITS1 and ITS2) (Ruiz, I.R.G., Silva, E.P. and Brison, O., unpublished results) and the intergenic spacer (IGS) accumulated several mutations in 2n and 4n and have little similarity to *X. laevis* spacers, except for the enhancers (Alvares, L.E., Brison, O. and Ruiz, I.R.G., unpublished results). On the other hand, there are no data in the *O. americanus* 2n/4n system concerning low repeated sequences, as the globin gene family.

Several  $\alpha$  and  $\beta$  globin genes with large introns are linked in *X. laevis*. They are organized in two  $\alpha\beta$ -globin clusters expressed in the same order as in the chromosomal cluster (Jeffreys *et al.*, 1980; Patient *et al.*, 1982; Hosbach *et al.*, 1983). In birds and mammals, the  $\alpha$  and  $\beta$  genes are arranged in separate clusters. In spite of the way the globin gene family is organized and expressed, all the component genes are structurally alike, composed by three exons and two introns, except for the pseudogenes often present in the  $\alpha\beta$  clusters (Breathnach and Chambon, 1981). Although there is a high variability in the sequence and size of the introns,

<sup>1</sup> Departamento de Biologia Geral, ICB, UFMG, Caixa Postal 486, 31270-901 Belo Horizonte, MG, Brasil. Send correspondence to M.D.P.A.

<sup>2</sup> UMR 103, Unité Mixte CNRS-bioMérieux, Lyon, France.

<sup>3</sup> Pfizer Inc., Groton, USA.

<sup>4</sup> Laboratório de Genética, Instituto Butantan, SP, Brasil.

their position is evolutionarily conserved in each gene. Some organisms, like Nematoda, have seven exons and six introns in a duplicated globin gene (Dixon *et al.*, 1992). An intron interrupts the exon II sequence in Nematoda, similar to plant globin genes (Brown *et al.*, 1984). The chironomid globin genes are intronless (Antoine and Niessing, 1984), although in *Chironomus thummi* two globin genes were recently found to have introns (Kao *et al.*, 1994).

In this paper, we describe the  $\alpha$ -globin major type gene from the *O. americanus* 2n and 4n globin clusters. A striking observation was the definite lack of intron 2 in both 2n and 4n amplified DNAs.

## MATERIAL AND METHODS

### Organisms

*O. americanus* 2n and 4n were collected in Botucatu and São Roque (São Paulo State, Brazil), and their ploidy level was confirmed by analysis of metaphase spreads.

### DNA preparation

Blood was extracted from the heart of *O. americanus* 2n and 4n by puncturing with a Pasteur pipette treated with 3.8% sodium citrate. DNA was extracted with proteinase K/phenol and treated with RNase, as described in Cortadas and Ruiz (1988).

### DNA amplification

The PCR reaction (Saiki *et al.*, 1985; Müllis and Faloona, 1987) was carried out with Taq DNA polymerase (The Perkin-Elmer Corporation, Norwalk, CT) and the buffer provided by the manufacturer in the presence of 200  $\mu$ M dNTP, 0.5  $\mu$ M of each primer, 50 ng of template DNA, and 2.5 units of enzyme in a final volume of 100  $\mu$ l for 30 cycles using a Perkin-Elmer Cetus thermalcycler. The first cycle of amplification consisted of 3-min denaturation at 95°C, 2-min annealing at 70°C, and 3-min extension at 72°C. The amplification products were size fractionated by agarose gel electrophoresis. The DNA primers (synthesized at the "Laboratoire Synthèse Nucléique" CNRS/bioMérieux, Lyon, France) were:

5'-CCCTCTAGAACGATATTTGGATGTCAGAAC-3'  
(antisense);

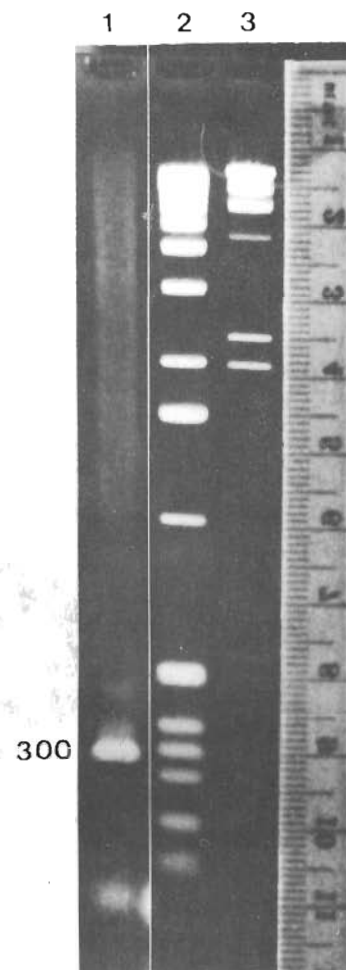
5'-CCGAATTCACCAAACTTACTTCCCT-3'  
(first sense) and

5'-CTGAGTGACCTCCATGC-3' (second sense).

In some experiments, restriction recognition sequences were incorporated at the 5' end of the first sense (*Eco*RI) and antisense (*Xba*I) primers to facilitate cloning (underlined).

### Cloning of PCR products

The PCR products were electrophoresed in 1.2% agarose gels in TBE buffer (89 mM Tris-base, 89 mM boric acid, 2 mM EDTA, pH 8.0) containing 0.5  $\mu$ g/ml ethidium bromide. The bands corresponding to the amplified products were cut out, and the DNA was recovered (Gene Clean II, Bio101 Inc.). The purified DNA was added to 10  $\mu$ l of 10x Polymerase I buffer (500 mM Tris-HCl, pH 7.5, 100 mM MgCl<sub>2</sub>, 10 mM DTT, 500  $\mu$ g/ml bovine serum albumin, 625  $\mu$ M dNTP), containing 1 mM rATP, 10 U T4 polynucleotide kinase and 10 U *E. coli* DNA polymerase I, Klenow fragment. Following incubation at 37°C for 1 hour, the DNA was



**Figure 1** - PCR amplification of  $\alpha$ -globin gene sequences from *O. americanus* 2n (specimen no. 1, 2n). The PCR product (insert approximately 300 bp, lane 1) was electrophoresed, eluted, cloned, and sequenced. The resulting pCD25 clone contains sequences of exons II and III, and lacks intron 2. Lane 2: BRL Gibco 1-kb ladder; lane 3:  $\lambda$  DNA/*Hind*III fragments.

precipitated with cold ethanol. Seventy-five nanogram of the resuspended fragments was ligated to 50 ng of *Sma*I-linearized pUC19 using standard conditions (Sambrook *et al.*, 1989). The ligation mixture was used to transform *E. coli* DH5 $\alpha$ , and the clones were screened with the probe pXG6C1 (Kay *et al.*, 1980), containing the  $\alpha$ -globin cDNA of *Xenopus laevis*, labelled with  $\alpha^{32}$ PdCTP. In another experiment, the PCR product was cloned as an *Eco*RI-*Xba*I fragment in pGEM3-Z (Promega, Madison, WI) to produce pCD25 (2n origin).

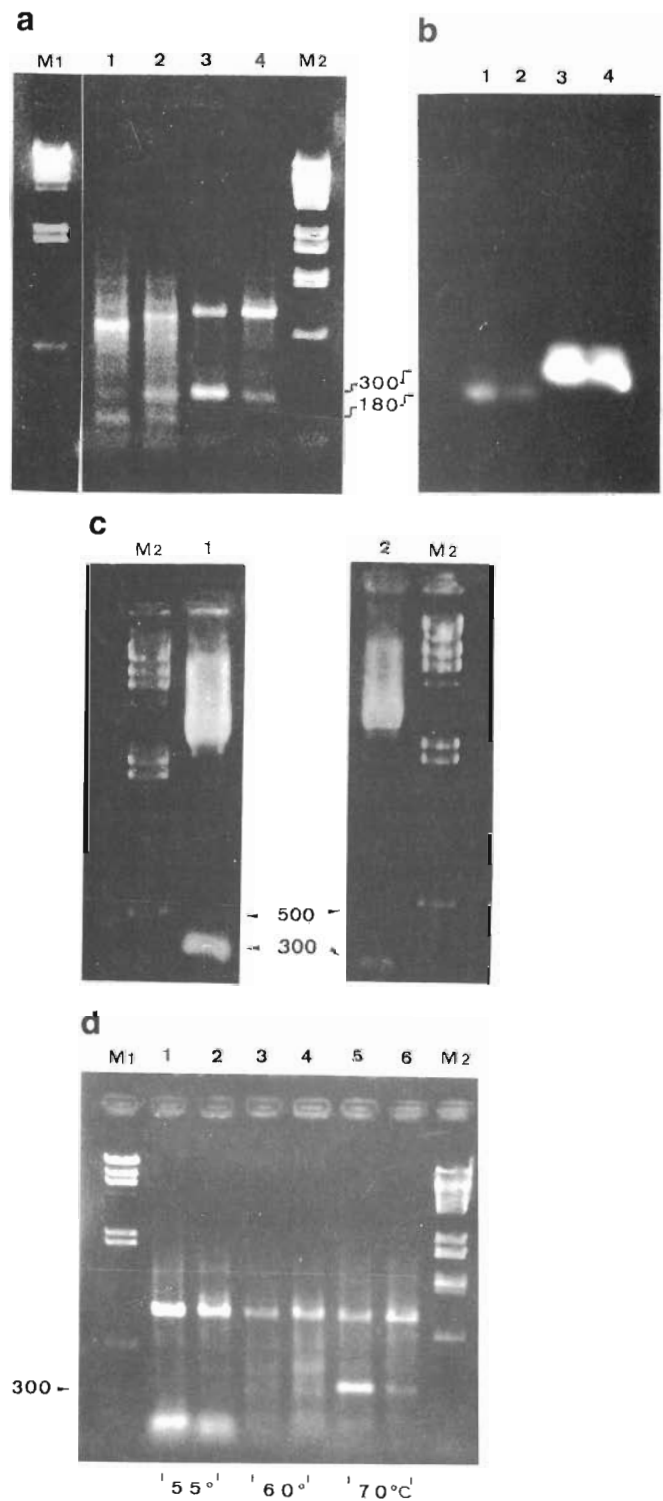
## Sequencing

The Sequenase Version 2 DNA sequencing kit (United States Biochemical) was employed to perform the sequencing (Sanger *et al.*, 1977) of selected inserts, according to the manufacturer's instructions. The primers used were those included in the kit, complementary to the flanking regions of the plasmid vector.

## RESULTS AND DISCUSSION

In the present work, the genes that code for the  $\alpha$ -globins of adult *Odontophrynus americanus* individuals were studied in an attempt to better understand the evolutionary differences between diploid and tetraploid specimens. The PCR technique was used to amplify globin sequences from genomic DNA obtained from blood cells. Two highly conserved exon regions were chosen, one of them located in the second exon (positions 114-131, first sense primer) and the other in the 3'-end of the third exon (positions 402-422, antisense primer). A third exon region was used as a second sense primer, downstream the exon II (positions 246-262).

Figure 1 shows the approximately 300-bp long amplified *Eco*RI/*Xba*I fragment of 2n origin, obtained with the first sense and antisense primers. Figure 2a shows another experiment, in which the PCR amplified fragments were approximately 180 bp and 300 bp long, and were obtained by combining either the second sense primer or the first sense primer with the antisense primer, respectively. These fragments were positive to the  $\alpha$ -globin probe (Figure 2b). Following elution and cloning of the PCR-amplified fragments, three clones were obtained: pCD25, of 2n origin, which contained a 321-bp insert including *Eco*RI and *Xba*I restriction recognition sequences (not shown); clones pOGL2.1 and pOGL4.12, of 2n and 4n origin, respectively, both inserts which contained 300 bp (Figure 2c). Analysis of the sequences obtained showed that all cloned fragments contained most of the exon II sequences and all of the exon III sequence. Intron 2 was completely lack-



**Figure 2** - PCR amplification of  $\alpha$ -globin gene sequences from *O. americanus* 2n (specimen no. 2, 2n) and 4n (specimen no. 1, 4n). a) Primer pairs used: second sense and antisense primers (lanes 1 and 2); first sense and antisense primers (lanes 3 and 4). b) Autoradiography of gel (a) hybridized to the pXG6C1 probe (*X. laevis*  $\alpha$ -globin cDNA). Due to their relative mobility, both positive 300-bp and 180-bp amplified fragments were not supposed to contain the intron 2. The PCR products (300 bp) were eluted from the gel and cloned. c) The selected clones pOGL2.1 (2n) and pOGL4.12 (4n) were digested with *Eco*RI and *Bam*HI to isolate the insert (lane 1: 2n; lane 2: 4n). Sequencing of the inserts confirmed the presence of exons II and III, and the lack of intron 2. d) The optimal annealing temperature for PCR amplification of globin sequences was 70°C (lanes 5 and 6, for 2n and 4n DNA, respectively). Markers were  $\lambda$ HindIII (M1) and  $\lambda$ BstI (M2).

```

First sense primer
114                               132
2n ACCAAACTT ACTTCCTAG TTTGACTTC CACCACAATT CAAACAGAT
4n .....
X1 .....

164
2n CAGTGTCAT GGCAAGAAAG TTGTGGATGC TCTGAATGAA GCTTCCAACC
4n .....
X1 .....

                               second sense primer
214                               246
2n ATTTGATAA CATCGCTGGA AGCATGAGCA AGCTGAGTGA CCTCCATGCC
4n .....
X1 .....

264 .                               (i)                               309
2n TATGACCTGA GAGTGGACCC TGGCAACTTC CCATTGCTGG CCCATGATAT
4n .....                               .....G.....
X1 .....                               .....A.....

314
2n ATTGGTGGTT GTTGCTATGA ACTTCCCTAA GCAGTTTGTAT CCTGCAACCC
4n .....
X1 .....

364                               402 antisense primer
2nATAAGGCCCT GGATAAGTTCTTGGCTACCGTATCTACTGTCTGACATCCAAATATCGT
4n .....
X1 .....

```

**Figure 3** - Inserts of *O. americanus* clones pCD25 (2n), pOGL2.1 (2n) and pOGL4.12 (4n) were sequenced and aligned to the corresponding sequences of the  $\alpha$ -globin major cDNA of *Xenopus laevis* (Kay et al., 1980). The first sense primer starts at the 114th position relative to the first nucleotide of exon I of the cDNA (corresponding to the 22nd position in exon II). The second sense primer starts at position 246. The antisense primer corresponds to the final sequences of the coding region of exon III, beginning at position 402, as indicated. Intron 2 (i) should be located between the positions 296/297. Except for the substitution at position 309 (transition G/A) and the absence of intron 2, the whole sequenced region is identical among *O. americanus* 2n, 4n and *X. laevis*.

ing (Figure 3). As these sequences were aligned with the major  $\alpha$ -globin sequence of *Xenopus laevis*, only one substitution (transition) was noted: A/G, corresponding to nucleotide 309 of the cDNA (Figure 3). This substitution altered amino acid number 104, which is an asparagine in *X. laevis*, to aspartic acid in *O. americanus*. This amino acid holds the G11 position in the  $\alpha$ -globin chain, inside the central cavity, between G and B helices. Although the G helix of  $\alpha$  and  $\beta$  chains are in contact with each other, the substitution to aspartic acid does not seem to alter the function of the hemoglobin molecule.

Surprisingly, in two independent experiments performed in two different laboratories, one experiment with 2n and the other with 2n and 4n DNA, the sequence analysis of the cloned fragments did not show the second intron, which is usually found in vertebrate

globin genes. The base pairs that would flank the introns remained intact. As the primer sense for PCR amplification was located in the second exon, it is not actually possible to know if the first intron was present or not.

Many kinds of intronless genes have been referred to in different eukaryotic species, e.g., the human gene coding for calmodulin-like protein (Koller and Strehler, 1993), and the glutathione S-transferase genes in *Drosophila melanogaster* (Toung et al., 1993). Moreover, absence of introns has been observed in the chironomid insect globin genes. These genes have all characteristics of productive genes and are expressed *in vivo* (Antoine and Niessing, 1984). Also, Vanin et al. (1980), Nishioka et al. (1980) and Lueders et al. (1982) reported mouse  $\alpha$ -globin pseudogenes that completely lack both globin intervening sequences. These pseudogenes may be incorporated into other chromosomal sites than the globin clusters, and do not code for functional peptides. The pseudogenes found outside the globin clusters are probably related to retrotransposition (Sharp, 1983), instead of gene duplication followed by deletions (gene conversion) (Little, 1982). Analysis of the flanking sequences of the globin gene may provide evidence of the pseudogene origin.

The amplification method used on the studied fragments should generate both kinds of globin genes, the productive one with its intron 2, and the putative pseudogene. However, electrophoretic analysis of the PCR products showed only one positive low molecular weight band (Figures 1 and 2b), that was eluted and cloned. Sequencing of these fragments proved them to belong to a globin-like sequence, lacking intron 2. The PCR amplification that produced the obtained sequences was not easily achieved: several assays were carried out in both laboratories under different experimental conditions, before any positive results could be obtained (Figure 2d). Did the PCR fail to amplify the "true"  $\alpha$ -globin gene due to some special characteristics of its introns? Or, is the intronless gene the only sequence related to the primers used? Cloning and analysis of genomic libraries, containing the whole globin cluster, might help answer these questions. PCR amplification of the 5'-end of the globin gene would be very helpful in discovering the existence of intron 1. Moreover, functional analysis at the gene and protein levels is required.

## ACKNOWLEDGMENTS

We are indebted to Dr. Nelson Leon for the valuable editorial revision of the manuscript, and to Luiz A.T. Dias for photography, both from Laboratório de Genética, Instituto Butantan, São Paulo, Brazil.

Publication supported by FAPESP.

## RESUMO

Os principais genes  $\alpha$ -globina dos anfíbios *Odontophrynus americanus* 2n e 4n foram estudados depois de clonagem e seqüenciamento de fragmentos obtidos através da técnica da PCR. Os primers utilizados para amplificação desses fragmentos foram escolhidos tendo como base seqüências conservadas publicadas de genes  $\alpha$ -globina de diferentes espécies. As seqüências estudadas continham o exon II e exon III do gene  $\alpha$ -globina. Foi observado que o intron 2 estava inteiramente ausente nos fragmentos amplificados tanto de origem 2n quanto nos de 4n. Quando as seqüências obtidas foram comparadas entre si e com a correspondente de *Xenopus laevis*, foi observado que são altamente conservadas. A possibilidade dessas seqüências serem pseudogenes é discutida.

## REFERENCES

- Antoine, M. and Niessing, J. (1984). Intronless globin genes in the insect *Chironomus thumi thumi*. *Nature* 310: 795-798.
- Beçak, M.L. and Beçak, W. (1974). Studies on polyploid amphibians, karyotype evolution and phylogeny of the genus *Odontophrynus*. *J. Herpetol.* 8: 337-341.
- Beçak, W. and Goissis, G. (1971). DNA and RNA content in diploid and tetraploid amphibians. *Experientia* 27: 345-346.
- Breathnach, R. and Chambon, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* 50: 349-383.
- Brown, G.G., Lee, J.S., Brisson, N. and Verma, D.P.S. (1984). The evolution of a plant globin gene family. *J. Mol. Evol.* 21: 19-32.
- Cortadas, J. and Ruiz, I.R.G. (1988). The organization of ribosomal genes in diploid and tetraploid species of the genus *Odontophrynus* (Amphibia, Anura). *Chromosoma* 96: 437-442.
- Dixon, B., Walker, B., Kimmins, W. and Pohajdak, B. (1992). A nematode hemoglobin gene contains an intron previously thought to be unique to plants. *J. Mol. Evol.* 35: 131-136.
- Hosbach, H.A., Wyler, T. and Weber, R. (1983). The *Xenopus laevis* globin gene family: chromosomal arrangement and gene structure. *Cell* 32: 45-53.
- Jeffreys, A.J., Wilson, V., Wood, D. and Simons, J.P. (1980). Linkage of adult  $\alpha$ - and  $\beta$ -globin genes in *X. laevis* and gene duplication by tetraploidization. *Cell* 21: 555-564.
- Kao, W.-Y., Trewitt, P.M. and Bergtrom, G. (1994). Intron containing globin genes in the insect *Chironomus thummi*. *J. Mol. Evol.* 38: 241-249.
- Kay, R.M., Harris, R., Patient, R.K. and Williams, J.G. (1980). Molecular cloning of cDNA sequences coding for the major  $\alpha$ - and  $\beta$ -globin polypeptides of adult *Xenopus laevis*. *Nucleic Acids Res.* 8: 2691-2707.
- Koller, M. and Strehler, E.E. (1993). Functional analysis of the promoters of the human CaMIII calmodulin gene and of the intronless gene coding for a calmodulin-like protein. *Biochim. Biophys. Acta* 1163: 1-9.
- Little, P.F.R. (1982). Globin pseudogenes. *Cell* 28: 683-684.
- Lueders, K., Leder, A., Leder, P. and Kuff, E. (1982). Association between a transposed  $\alpha$ -globin pseudogene and retrovirus-like elements in the BALB/c mouse genome. *Nature* 295: 426-428.
- Müllis, K.B. and Faloona, F.A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. In: *Methods in Enzymology* (Wu, R., ed.). Vol. 3. Academic Press, San Diego, CA, pp. 335-350.
- Nishioka, Y., Leder, A. and Leder, P. (1980). Unusual  $\alpha$ -globin-like gene that has clearly lost both globin intervening sequences. *Proc. Natl. Acad. Sci.* 77: 2806-2809.
- Patient, R.K., Banville, D., Brewer, A.C., Elkington, J.A., Greaves, D.R., Lloyd, M.M. and Williams, J.G. (1982). The organization of the tadpole and adult  $\alpha$ -globin genes of *Xenopus laevis*. *Nucleic Acids Res.* 10: 7935-7945.
- Ruiz, I.R.G. and Brison, O. (1989). Methylation of ribosomal cistrons in diploid and tetraploid *Odontophrynus americanus* (Amphibia, Anura). *Chromosoma* 98: 86-92.
- Ruiz, I.R.G., Soma, M. and Beçak, W. (1981). Nucleolar organizer regions and constitutive heterochromatin in polyploid species of the genus *Odontophrynus* (Amphibia, Anura). *Cytogenet. Cell. Genet.* 29: 84-98.
- Saiki, R.K., Scharf, S., Faloona, F., Müllis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985). Enzymatic amplification of  $\beta$ -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230: 1350-1354.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74: 5463-5467.
- Sharp, P.A. (1983). Conversion of RNA to DNA in mammals: Alu-like elements and pseudogenes. *Nature* 301: 471-472.
- Toung, Y.-P., Hsieh, T. and Tu, C.-P. (1993). The glutathione S-transferase D genes. A divergently organized, intronless gene family in *Drosophila melanogaster*. *J. Biol. Chem.* 268: 9737-9746.
- Vanin, E.F., Goldberg, G.I., Tucker, P.W. and Smithies, O. (1980). A mouse  $\alpha$ -globin related pseudogene lacking intervening sequences. *Nature* 286: 222-226.

(Received March 6, 1997)