Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 6691–6695, September 1989
Genetics

A strategy to detect and isolate an intron-containing gene in the presence of multiple processed pseudogenes

(ribosomal protein genes/polymerase chain reaction/multigene family/cloning strategy)

BRENDAN DAVIES, SALVATORE FEO*, EDITH HEARD, AND MIKE FRIED

Department of Eukaryotic Genome Organization and Expression, Imperial Cancer Research Fund, P.O. Box 123, Lincoln's Inn Fields, London WC2A 3PX, United Kingdom

Communicated by W. F. Bodmer, June 2, 1989

ABSTRACT **We have devised a strategy that utilizes the polymerase chain reaction (PCR) for the detection and isolation of intron-containing genes in the presence of an abundance of processed pseudogenes. The method depends on the genomic DNA sequence between the PCR primers spanning at least one intron in the gene of interest, resulting in the generation of a larger intron-containing PCR product in addition to the smaller PCR product amplified from the intronless pseudogenes. A unique intron probe isolated from the larger PCR product is used for the detection of intron-containing clones from recombinant DNA libraries that also contain pseudogene clones. This method has been used successfully for the selective isolation of an intron-containing rat L19 ribosomal protein gene in the presence of multiple pseudogenes. Analysis of a number of mammalian ribosomal protein multigene families by PCR indicates that they all contain only a single gene with introns.**

Many highly expressed mammalian "housekeeping genes" are members of multigene families that contain multiple processed pseudogenes (1, 2). This complication has impeded the isolation of functional intron-containing genes for a number of these mammalian housekeeping genes, including many ribosomal protein genes. The mammalian ribosomal protein pseudogenes have all the hallmarks of being formed by reverse transcription as they lack introns, contain a poly(A) stretch at their 3′ ends, and have direct sequence repeats present at the 5′ and 3′ ends of the gene (3–7). For each ribosomal protein there is thought to be only one active intron-containing gene that is usually difficult to clone in the presence of multiple pseudogenes (8–30 copies). Thus although many ribosomal protein cDNAs have been isolated (8–29), only a few functional intron-containing ribosomal protein genes have been cloned (3–5, 7, 30–32) as the different isolation strategies used involve laborious cloning and screening procedures. In this communication we describe a simple method utilizing the polymerase chain reaction (PCR) (33) to detect and clone intron-containing genes that are members of multigene families containing many processed pseudogenes. Using this method, we have selectively isolated a clone containing a region of the intron-containing rat L19 ribosomal protein gene. We also present evidence to suggest that a number of the mammalian ribosomal protein multigene families contain only one gene with introns.

MATERIALS AND METHODS

PCRs were performed on genomic DNA (1–2 μg) in 100 μl containing 50 mM KCl, 10 mM $MgCl_2$, 0.01% gelatin, 10% dimethyl sulfoxide, 200 μM (each) dATP, dCTP, dGTP, and TPP, 1 μM each primer, and 2.5 units of Taq polymerase (*Thermus aquaticus* DNA polymerase) (Perkin–Elmer/Cetus). All PCRs were carried out in a programable Dri-Block (Techne) and consisted of 30 cycles of 2 min at 94°C (denaturation); 2 min at 59°C (annealing); 9.9 min at 70°C (elongation). In the case of ribosomal protein L7a the annealing temperature was 63°C. The PCR products were analyzed by agarose gel electrophoresis on 1.2–1.4% gels and visualized with UV light following ethidium bromide staining. The DNA sequence was determined from either single- or double-stranded subclones or directly from the PCR product (in which case the required band was first excised from an agarose gel and electroeluted) using the Sequenase kit and protocols (United States Biochemical). For Southern transfer, 10 μg of rat genomic DNA was digested with a variety of restriction enzymes and fractionated on a 0.8% agarose gel before being blotted onto Hybond-N (Amersham) and hybridized with the relevant probes according to the manufacturer's recommendations. A rat *Mbo* I partial genomic library was constructed in λFIX (Stratagene) and packaged with Gigapack Gold (Stratagene) following the supplier's conditions. Partial digestion conditions were adjusted to reduce the bulk of the DNA to fragments of ≈15–20 kilobases (kb). Plaques (1.5×10^5) were screened as described (34) using a unique intron probe.

RESULTS

Strategy for Detecting and Cloning Intron-Containing Genes. The strategy used for the isolation of intron-containing genes in the presence of multiple intronless pseudogenes is presented in Fig. 1. Sets of short oligonucleotides [20–30 base pairs (bp)] derived from a previously sequenced cDNA are used as primers with genomic DNA in a PCR. If, in the intron-containing gene, an intron is located between the sequences homologous to the primers, the amplified DNA derived from this gene will be larger than the DNA amplified from the intronless pseudogenes by the size of the intron(s). This larger PCR product can be detected, separated, and isolated by gel electrophoresis (Fig. 1). Unique intron probes can be used to isolate intron-containing gene clones from genomic DNA libraries.

Isolation of a Rat L19 Ribosomal Protein Intron-Containing Gene. The application of this technique to the isolation of the intron-containing rat L19 ribosomal protein gene in the presence of its many processed pseudogenes is presented in Fig. 2. One oligonucleotide derived from the middle and another derived from the 3′ region of the previously described rat L19 cDNA sequence (12) were synthesized and used as primers with total rat cellular DNA in a PCR (see *Materials and Methods*). Agarose gel electrophoresis of the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: PCR, polymerase chain reaction.

*Permanent address: Istituto di Biologia dello Sviluppo, Consiglio Nazionale delle Ricerce, Via Archirafi, 20, 90123 Palermo, Italy

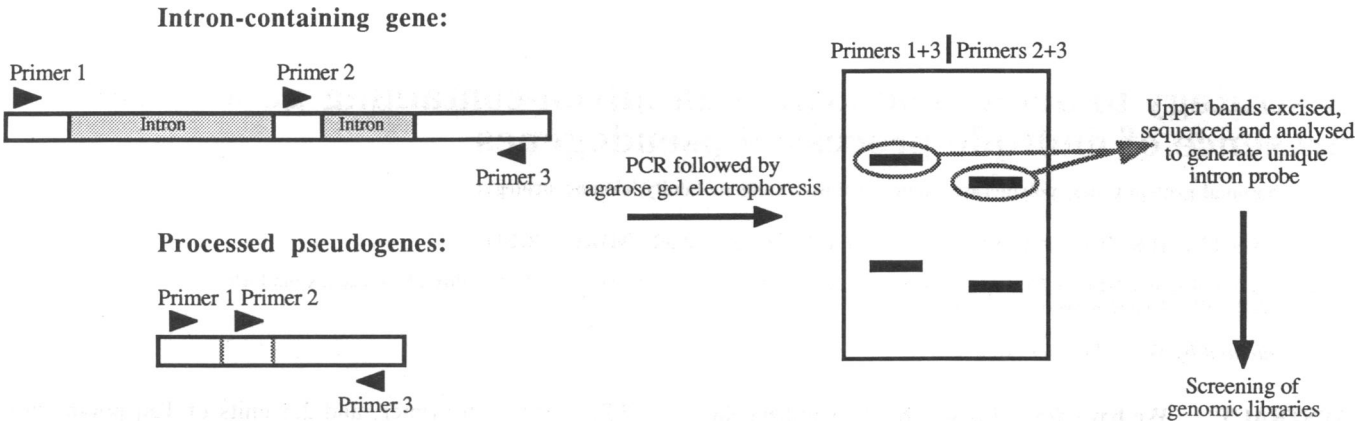


FIG. 1. Strategy for the detection and isolation of intron-containing genes in the presence of multiple processed pseudogenes. The single genomic intron-containing gene is shown above one of the multiple processed pseudogenes. The introns are shown as stippled boxes and the exons are shown as open boxes. Pairs of oppositely oriented exon-specific oligonucleotide primers of 20–30 bp (primers 1 and 3 or 2 and 3), chosen from the cDNA sequence, are used in a PCR with genomic DNA and the products are analyzed by agarose gel electrophoresis. Each pair of primers will generate a band, corresponding to their distance apart in the processed pseudogenes, which will be of the same size as that predicted by their distance apart in the cDNA. Additionally, if, in the intron-containing gene, any of these pairs of primers happens to be separated by one or more introns, a larger band will be observed. These larger bands can be excised and analyzed by sequencing and Southern blotting to derive a unique intron probe for the screening of genomic libraries.

PCR products revealed two amplified bands (Fig. 2*A*). The size of the lower band (335 bp) was that expected for the amplification of L19 processed pseudogenes and this identity was confirmed by DNA sequence analysis of the product, which showed both high sequence homology and a structure identical to the L19 cDNA (i.e., no introns) (data not shown). The upper band of 996 bp was excised from the gel and the sequence of the 410 bp at its 5′ end was determined (Fig. 2*B*). The first 26 bp of this sequence is identical to the exon sequence of the L19 cDNA (nucleotides 394–419) and is followed by 305 bp of unidentifiable sequence before rejoining the cDNA exon sequence at nucleotide 420 and continuing for at least another 76 bp. This structure is what would be expected for the presence of an intron in the genomic DNA separating an exon ending at cDNA nucleotide 419 and an exon starting at cDNA nucleotide 420.

To determine whether the putative intron sequence was unique and could be used for cloning the intron-containing L19 ribosomal protein gene, a 260-bp probe extending from the 5′ end of the upper PCR product to the first of the two *Nco* I sites (Fig. 2*B*) was used to probe Southern blots of rat cellular DNA. This probe, which contains 25 bp of exon sequence, detected only a single band when rat DNA was digested with *Eco*RI, *Bam*HI, *Hin*dIII, *Bgl* II, or *Pst* I restriction enzyme (Fig. 3*A*), indicating that only a single copy of the intron-containing gene is present in the rat genome. This is in contrast to the multiple bands detected when a similar blot of restricted DNA is hybridized with a probe derived from the lower PCR band (Fig. 3*B*), which also detects the many L19 processed pseudogenes that are widely distributed in the rat genome.

The unique intron-containing probe was used to screen an *Mbo* I partial library of rat cellular DNA cloned in Lambda Fix (Stratagene). One positive recombinant clone (λRL19.1), containing an insert of 10.5 kb (Fig. 4*A*), was detected out of 1.5×10^5 plaques screened. To locate the region of λRL19.1

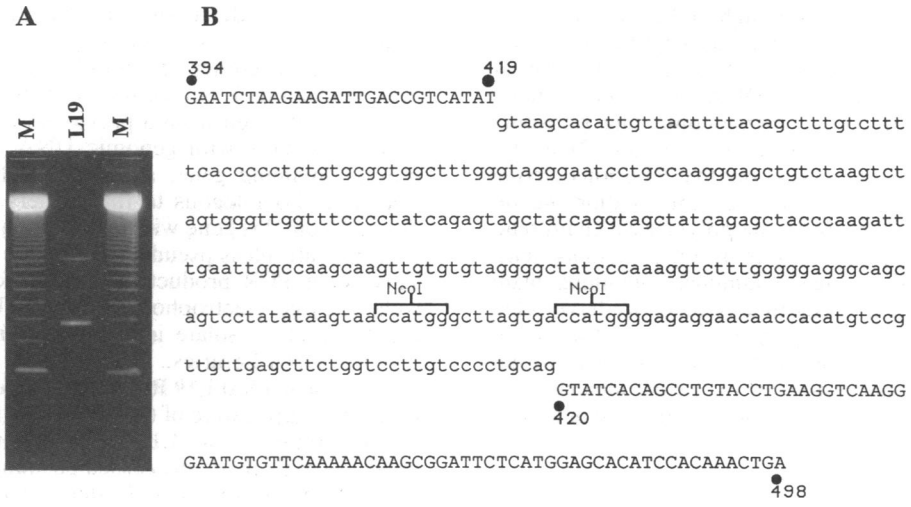


FIG. 2. Analysis of the PCR products generated from the rat L19 ribosomal protein genes. (*A*) Agarose gel electrophoresis of the PCR products generated from rat genomic DNA using primers derived from the rat L19 cDNA sequence (12). One primer corresponds to L19 cDNA nucleotides 394–420 and the other to the complement of L19 cDNA nucleotides 689–715. Two PCR products are detected in the L19 track corresponding to the intron-containing gene (996-bp upper band) and the processed pseudogenes (235-bp lower band). The tracks labeled M each contain a 123-bp ladder (BRL) as size markers. (*B*) Partial DNA nucleotide sequence of the putative intron-containing gene. Four-hundred ten base pairs of the DNA sequence of the upper L19 PCR product (see *A*) was determined and is shown. Uppercase characters denote homology to the cDNA sequence and lowercase characters denote putative intron sequence. The numbering is that of the cDNA. The two *Nco* I sites, referred to in the text, are indicated.

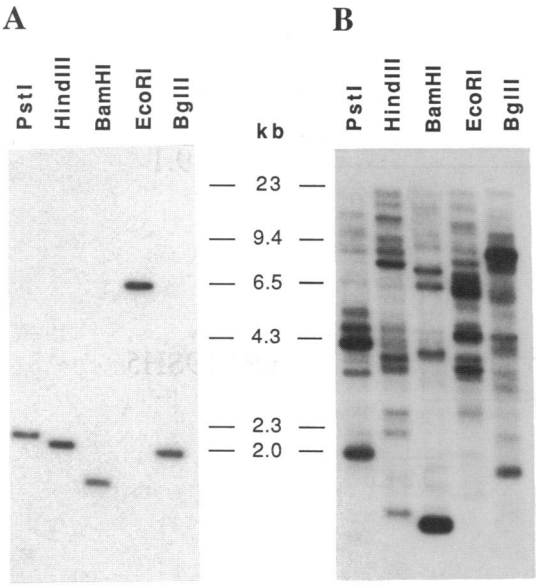


FIG. 3. Southern blot analysis of rat genomic DNA using L19 intron- and exon-specific probes. (*A* and *B*) Identical Southern blots probed either with an intron-specifc probe (*A*) or an exon-specific probe (*B*). The intron-specific probe was the 260-bp fragment derived from the upper PCR band (Fig. 2*A*) extending from the 5′ end of the upstream primer to the first of the two *Nco* I sites (Fig. 2*B*). The exon-specific probe was the lower PCR band (Fig. 2*A*). Positions of the λ *Hin*dIII size markers are indicated.

containing the L19 sequence the three insert fragments generated by restriction of cloned DNA with *Sal* I and *Hin*dIII were tested by PCR for the presence of L19-specific amplification products. Only the 1.9-kb fragment extending from the *Sal* I site in the Lambda Fix polylinker to a *Hin*dIII site within the insert was found to be positive by PCR and only for the L19 larger 996-bp product (see Fig. 2*A*). This 1.9-kb *Sal* I–*Hin*dIII fragment was subcloned into Bluescript plasmid (pRL19SH5) (Fig. 4*A*) and the 1.2 kb of sequence between the *Sal* I site and *Xba* I site was determined and compared to the L19 cDNA sequence (12). The results of the analysis indicate that the λRL19.1 clone contains the 3′ half of the L19 intron-containing gene (Fig. 4). The genomic sequence starts at cDNA nucleotide 363 and extends to cDNA nucleotide 419, where it reaches the 305-bp intron (intron a in Fig. 4) previously detected by PCR (Fig. 2). The sequence then continues between cDNA nucleotides 420 and 530, at which point a second intron of 357 bp (intron b in Fig. 4) is found. The sequence then rejoins the cDNA sequence at nucleotide 531 and continues past the start of the poly(A) tail found in the cDNA sequence at nucleotide 736 (Fig. 4*C*). There is near perfect concordance of the sequences at the two intron/exon joins with the consensus sequences for splice donor and acceptor sites (Fig. 4*B*) (35). There are two differences between the exon sequence of the cloned genomic DNA sequence and that of the cDNA sequence (12), which both occur in the 3′ untranslated region. One is a base change at nucleotide 683 (thymine to cytosine) and the other is an additional thymine in the genomic DNA after nucleotide 694 (Fig. 4*C*). These could be due to cloning artifacts, sequencing errors, or DNA polymorphisms between the different rat strains used.

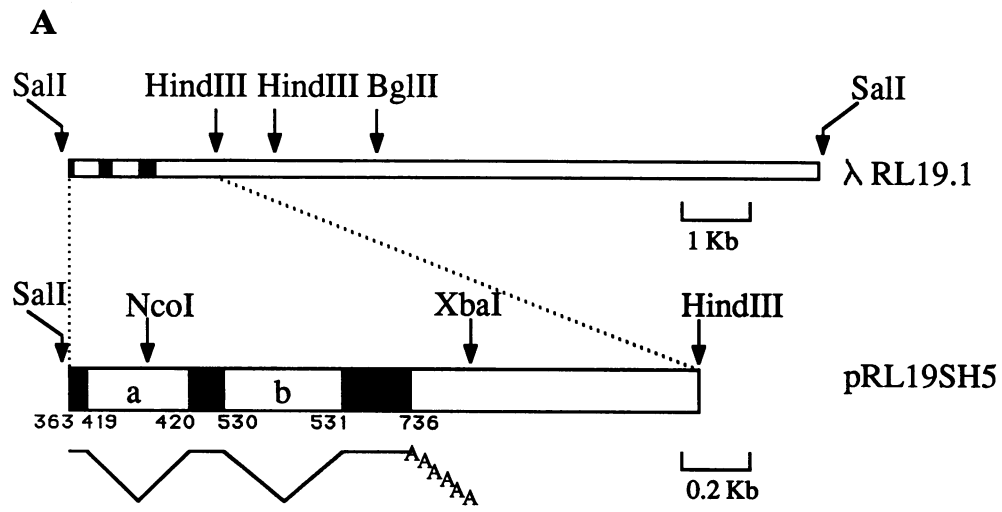
Detection of Other Ribosomal Protein Intron-Containing Genes. By using oligonucleotide primers derived from a number of different mammalian ribosomal protein cDNA sequences we were able to identify a larger PCR product in addition to the lower pseudogene band after PCR analysis for each gene tested (Fig. 5). These results indicate that this technique will be useful for the cloning of the intron-containing genes of other pseudogene-containing ribosomal protein multigene families. In most cases only a single predominant larger PCR product is detected.

DISCUSSION

We have used the PCR to detect and isolate a single-copy intron-containing gene in the presence of an abundance of processed pseudogenes. The method depends on the sequence between the primers used in the PCR spanning at least one intron in the gene of interest and thus generating a larger amplified fragment than that generated by PCR from the intronless pseudogenes (Fig. 1). Unique intron probes derived from the larger PCR product can then be used to isolate intron-containing gene clones selectively from genomic recombinant DNA libraries that also contain multiple pseudogene clones.

Ribosomal protein genes are of great interest as they are a set of genes whose expression is coordinately regulated in response to the varying protein synthesis requirements of the cell. It has not been easy to analyze the functional intron-containing mammalian ribosomal protein genes, as a result of the difficulty of isolating these single (or low) copy genes in the presence of multiple processed pseudogenes, using conventional cloning strategies (6). At present only six mammalian intron-containing ribosomal protein genes have been isolated, one by genetic complementation (30), four by laborious cloning and screening (3–5, 31), and one fortuitously (7, 32). The strategy described above has been successfully used to isolate part of the intron-containing gene for the rat ribosomal protein L19. Comparison of the sequence of the larger L19 PCR product (Fib. 2*B*) and the genomic clone (Fig. 4) with the L19 cDNA sequence (12) has revealed the presence of *bona fide* introns located between consensus splice donor and acceptor sites (Fig. 4*B*). In addition, the L19 intron-containing gene does not contain a poly(A) stretch in the position where the poly(A) tail is located in the L19 cDNA (Fig. 4*C*) (12), which is a feature of ribosomal protein pseudogenes and other pseudogenes formed by retroposition (1, 2). Thus the larger L19 PCR product (Fig. 2*A*) appears to represent a true intron-containing gene and does not result from non-intron sequence generated by rearrangement within a processed pseudogene or from transposition of another DNA sequence into an L19 pseudogene sequence.

In previous analyses (3–5, 7, 31) it has not been possible to determine whether more than one intron-containing gene is present, in addition to the multiple pseudogenes, in ribosomal protein multigene families. Our results strongly suggest that there is only one rat L19 ribosomal protein gene that contains introns as only one larger PCR band was detected with each of two sets of primers derived from different regions of the L19 cDNA (Fig. 2*A*, Fig. 5). In addition only one sequence was derived from direct sequencing of the larger PCR product (Fig. 2*A*) and only bands consistent with one single copy gene were found when this larger PCR product was used as a probe on rat genomic DNA (Fig. 3*A*). We cannot rule out the existence of another L19 ribosomal protein gene that contains a very large intron that would be poorly amplified in the PCR because of the great distance between the primers or the possibility that another L19 gene exists that contains a sequence within an intron that is poorly traversed by the polymerase during the PCR. Furthermore, there may be another L19 ribosomal protein gene in which an intron is found in the middle of one of the oligonucleotide sequences used as a primer for the PCR and thus would not be able to function efficiently for PCR. In the experiments described here this is less likely as the same introns were identified using two different sets of oligonucleotides as primers. It is interesting to point out that the two primers chosen for the PCR were not entirely homologous to continuous L19 exon DNA sequence. An intron is present in the L19 genomic



B

	Splice Donor	Intron Size	Splice Acceptor
Intron a	TAT/GTAAGC	- - - -(305 bp)- - -	-GTCCTTGTCCCCTGCAG/GTA
Intron b	GGC/GTAAGT	- - - -(357 bp)- - -	-TCTTTTCCTTATACCAG/TGA
Consensus	$^{C}_{A}AG/GT^{A}_{G}AGT$		$(^{T}_{C})_{11}N^{T}_{C}AG/G$

C

676

L19cDNA TAGTGGCTAGGCTATGGCC CACATGGATCAGTCATTAAAATAAAACAAGCC
pRL19SH5 TAGTGGCCAGGCTATGGCCTCACATGGATCAGTCATTAAAATAAAACAAGCC

* *

736

L19cDNA TTCGTCCTTG(A)$_n$
pRL19SH5 TTCGTCCTTGTCCATTGCCCTCTTTGTAGCAGTTTTGGCTGGTGTGTCTTAC

FIG. 4. Analysis of the intron-containing rat L19 ribosomal protein gene. (*A*) Maps of the λ clone (λRL19.1) and a plasmid subclone, pRL19SH5, containing the 3′ region of the L19 intron-containing gene. The *Sal* I sites are from the λFIX polylinker. The black shading represents exon sequences and the numbering is that of the L19 cDNA (12). The two introns (a and b) within the coding sequence are indicated. The predicted mRNA for this region of the gene is shown diagrammatically beneath the map of the plasmid subclone. (*B*) Sequence comparisons of splice donor and acceptor sites. The DNA sequences around the splice donor and acceptor sites of introns a and b (see *A*) are presented above the mammalian consensus sequences (35). The sizes of the introns are also indicated. (*C*) Part of the DNA sequence at the 3′ end of the L19 intron-containing gene is shown beneath the sequence of the L19 cDNA (12). The putative AATAAA poly(A) signal is shown boxed and the sequence complementary to the 3′ PCR primer (see Fig. 2) is overlined. The two disparities between the published cDNA sequence (12) and that of the genomic clone are indicated by asterisks. The site of the poly(A) tail in the cDNA is indicated by $(A)_n$.

DNA between the penultimate nucleotide (thymine) and the last nucleotide (guanine) in the most 5′ primer, but this did not inhibit the PCR as fortuitously the first intron nucleotide is also a guanine so that the primer is completely homologous to the L19 genomic DNA over the exon/intron boundary (Fig. 2*A*). The sequence of the 3′ untranslated L19 genomic sequence was found to differ from the cDNA in the region of the middle of the 3′ primer by containing an additional nucleotide after nucleotide 694 (Fig. 4*C*), which also did not affect the PCR.

Using the arguments and caveats described above we feel that there is only one intron-containing gene for each of the other mammalian ribosomal protein genes analyzed as only one predominant larger PCR product is detected for each gene (Fig. 5). In the case of mouse ribosomal protein L7a, whose intron-containing gene has been cloned and analyzed, the larger PCR band is exactly the size expected for the correctly primed L7a fragment containing intron 1 (Fig. 5). In a similar analysis with mouse ribosomal protein genes L30 and L32 we have, in each case, detected only a single larger PCR band of the size predicted from the known intron sizes of these genes (data not shown). Thus for every mammalian ribosomal gene family analyzed by PCR we find evidence for only one intron-containing gene.

In this study we have shown that we can detect and isolate intron-containing genes in the presence of multiple processed pseudogenes. The method described here could also prove useful in mapping the location of introns in single copy genes

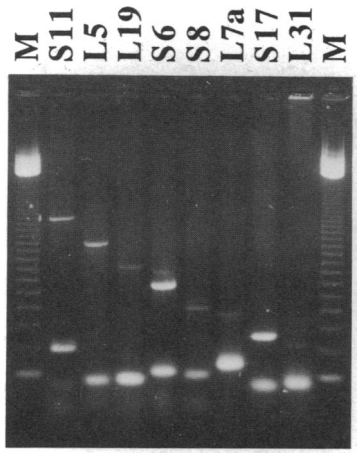


FIG. 5. PCR analysis of other mammalian ribosomal protein genes. Agarose gel electrophoresis of PCR products generated from rat, mouse, or human genomic DNA using oligonucleotide primers derived from cDNAs of different ribosomal proteins. For each ribosomal protein listed P1 indicates the cDNA nucleotide numbers used for the first PCR primer and P2 indicates the cDNA nucleotides from whose complement the second PCR primer was derived. Rat L5: P1, nucleotides (nt) 54–83; P2, nt 97–123 (9). Rat L19: P1, nt 36–60; P2, nt 90–111 (12). Rat S8: P1, nt 20–42; P2, nt 89–115 (24). Rat L31: P1, nt −21 to −3; P2, nt 18–45 (16). Rat S17: P1, nt 1–21; P2, nt 28–41 (15). Human S11: P1, nt 330–348; P2, nt 515–540 (26). Mouse L7a: P1, nt 1–21; P2, nt 28–41 (7). The tracks labeled M contain the 123-bp ladder size markers (BRL).

where only the cDNA has been isolated. We have isolated the 3′ region of the rat L19 ribosomal protein gene that contains two introns and have undertaken preliminary analyses of the 5′ region of the L19 gene that contains another three or four introns. In addition we have isolated clones of a number of the intron-containing ribosomal protein genes analyzed in Fig. 5. The strategy and methods used here should be applicable to the intron-containing genes of other processed pseudogene-containing multigene families. By this means we hope to be able to study the functional members of these gene families.

We thank Mr. Iain Goldsmith for his invaluable help in synthesizing the many oligonucleotide primers used in this work. We are also grateful to Drs. Clive Dickson, Nancy Hogg, and Clare Huxley for their helpful comments in the preparation of this manuscript.

1. Vanin, E. F. (1985) *Annu. Rev. Genet.* **19**, 253–272.
2. Weiner, A. M., Deininger, P. L. & Efstratiadis, A. (1986) *Annu. Rev. Biochem.* **55**, 631–661.
3. Wiedemann, L. M. & Perry, R. P. (1984) *Mol. Cell. Biol.* **4**, 2518–2528.
4. Dudov, K. P. & Perry, R. P. (1984) *Cell* **37**, 457–468.
5. Wagner, M. & Perry, R. P. (1985) *Mol. Cell. Biol.* **5**, 3560–3576.
6. Kuzumaki, T., Tanaka, T., Ishikawa, K. & Ogata, K. (1987) *Biochim. Biophys. Acta* **909**, 99–106.
7. Huxley, C., Williams T. & Fried, M. (1988) *Mol. Cell. Biol.* **8**, 3898–3905.
8. Chan, Y.-L., Lin, A., McNally, J. & Wool, I. G. (1987) *J. Biol. Chem.* **262**, 12879–12886.
9. Tamura, S., Kuwano, Y., Nakayama, T., Tanaka, S., Tanaka, T. & Ogata, K. (1987) *Eur. J. Biochem.* **168**, 83–87.
10. Lin, A., Chan, Y.-L., McNally, J., Peleg, D., Meyuhas, O. & Wool, I. G. (1987) *J. Biol. Chem.* **262**, 12665–12671.
11. Devi, K. R. G., Chan, Y.-L. & Wool, I. G. (1988) *DNA* **7**, 157–162.
12. Chan, Y.-L., Lin, A., McNally, J., Peleg, D., Meyuhas, O. & Wool, I. G. (1987) *J. Biol. Chem.* **262**, 1111–1115.
13. Tanaka, T., Kuwano, Y., Ishikawa, K. & Ogata, K. (1988) *Eur. J. Biochem.* **173**, 53–56.
14. Belhumeur, P., Paterne, G. D., Boileau, G., Claverie, J.-M. & Skup, D. (1987) *Nucleic Acids Res.* **15**, 1019–1029.
15. Nakanishi, O., Oyanagi, M., Kuwano, Y., Tanaka, T., Nakayama, T., Mitsui, H., Nabeshima, Y. & Ogata, K. (1985) *Gene* **35**, 289–296.
16. Tanaka, T., Kuwano, Y., Kuzumaki, T., Ishikawa, K. & Ogata, K. (1987) *Eur. J. Biochem.* **162**, 45–48.
17. Rajchel, A., Chan, Y.-L. & Wool, I. G. (1988) *Nucleic Acids Res.* **16**, 2347.
18. Tanaka, T., Wakasugi, K., Kuwano, Y., Ishikawa, K. & Ogata, K. (1986) *Eur. J. Biochem.* **154**, 523–527.
19. Gallagher, M. J., Chan, Y.-L., Lin, A. & Wool, I. G. (1988) *DNA* **7**, 269–273.
20. Davies, M. S., Henney, A., Ward, W. H. J. & Craig, R. K. (1986) *Gene* **45**, 183–191.
21. Chan, Y.-L. & Wool, I. G. (1988) *J. Biol. Chem.* **263**, 2891–2896.
22. Heinze, H., Arnold, H. H., Fischer, D. & Kruppa, J. (1988) *J. Biol. Chem.* **263**, 4139–4144.
23. Lott, J. B. & Mackie, G. A. (1988) *Gene* **65**, 31–39.
24. Chan, Y.-L., Lin, A., Paz, V. & Wool, I. G. (1987) *Nucleic Acids Res.* **15**, 9451–9459.
25. Lott, J. B. & Mackie, G. A. (1988) *Nucleic Acids Res.* **16**, 1205.
26. Tanaka, T., Kuwano, Y., Ishikawa, K. & Ogata, K. (1985) *J. Biol. Chem.* **260**, 6329–6333.
27. Lin, A., Chan, Y.-L., Jones, R. & Wool, I. G. (1987) *J. Biol. Chem.* **262**, 14343–14351.
28. Rhoads, D. D. & Roufa, D. J. (1985) *Mol. Cell. Biol.* **5**, 1655–1659.
29. Kuwano, Y., Nakanishi, O., Nabeshima, Y., Tanaka, T. & Ogata, K. (1985) *J. Biochem.* **97**, 983–992.
30. Rhoads, D. D., Dixit, A. & Roufa, D. J. (1986) *Mol. Cell. Biol.* **6**, 2774–2783.
31. Chen, I.-T. & Roufa, D. J. (1988) *Gene* **70**, 107–116.
32. Giallongo, A., Yon, J. & Fried, M. (1989) *Mol. Cell. Biol.* **9**, 224–231.
33. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science* **230**, 1350–1354.
34. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
35. Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472.