Volume 10 Number 22 1982 Nucleic Acids Research

Nucleotide sequence of a segment of human DNA containing the three tRNA genes

Kenneth L.Roy*, Howard Cooke+ and Richard Buckland†

*Department of Microbiology, University of Alberta, Edmonton, Canada T6G 2E9, +MRC Mammalian Genome Unit, Department of Zoology, West Mains Rd., Edinburgh EH9 3JT, and †MRC Clinical and Population Cytogenetics Unit, Western General Hospital, Edinburgh EH4 2XU, UK

Received 2 July 1982; Revised and Accepted 14 October 1982

ABSTRACT

A 1.65 kb segment of DNA from a human-λ recombinant, selected for human tRNA genes, was subcloned in the plasmid pAT 153 for sequence determination. Three human tRNA genes were found; one, a $tRNA^{lys}$ gene which appears to specify a tRNA identical in sequence to the previously described $tRNA_3^{lys}$ from rat liver (AAA,G); another a $tRNA^{gln}$ gene, the product of which would be expected to recognize only the codon CAG; and a third which would direct synthesis of a $tRNA^{leu}$ specific for CUA and G. Intervening sequences are not found in any of these genes. The three tRNA genes are separated by about 0.5 kb segments of DNA containing no apparent informational sequences. Examination of flanking sequences shows some similarities at the 5'-ends of the three genes, but less similarity to 5' flanking sequences of tRNA genes in other eucaryotes. All three tRNA genes direct synthesis of appropriate sized products in a HeLa cell lysate *in vitro* transcription system.

INTRODUCTION

The sequences of a number of eucaryotic tRNA genes have been reported, (1) but few of these have been determined for vertebrates. Müller and Clarkson (2) reported the sequence of the first 593 bases of the 3.18 kb repeated unit from *Xenopus laevis*, which includes genes for a $tRNA^{phe}$ and a $tRNA^{tyr}$. This repeated unit also contains genes for two $tRNA^{met}$ species (3) as well as $tRNA^{asp}$, $tRNA^{ala}$, $tRNA^{lys}$ and $tRNA^{leu}$ (Müller and Clarkson, unpublished, cited in Hofstetter *et al.*, (4), and Galli *et al.*, (5). More recently Santos and Zasloff (6) determined the sequence of two $tRNA^{met}$ genes of human origin, and Sekiya *et al.* (7) cloned and sequenced a 2 kb segment of rat DNA containing genes for a $tRNA^{asp}$, a $tRNA^{gly}$ and a $tRNA^{glu}$.

Although there is good evidence that tRNA genes are highly conserved (1,8) there is less information concerning conservation of their flanking sequences, and the occurrence of intervening sequences. Furthermore, although evidence is accumulating which suggests little conservation of tRNA gene arrangement in vertebrates (2,6,7,9) there is a need for more

0305-1048/82/1022-7313$ 2.00/0

extensive information to define the limits of this conservation.

In an investigation of human tRNA genes (R. Buckland, H. Cooke, K.L. Roy, J. Dahlberg and E. Lund, in preparation) a human-λ recombinant (λHt 4) was isolated which had several tRNA genes. Because of uncertainties in the determination of which tRNA genes were present in this recombinant, a large portion of the human insert was subcloned in a series of plasmids for sequence determination. We wish to describe the sequence of 1.65 kb of human DNA which contains three tRNA genes, all of which appear to be functional.

MATERIALS AND METHODS

Enzymes and Reagents

Restriction endonucleases were purchased from Bethesda Research Laboratories, New England Biolabs and Boehringer-Mannheim. The Klenow fragment of *E. coli* DNA polymerase I was from New England Nuclear and PL Biochemicals.

All $\alpha[^{32}P]$-deoxyribonucleoside triphosphates (at >600 Ci/mmol) were from New England Nuclear, as was the HeLa cell lysate *in vitro* transcription system. DEAE-Sephacel was a product of Pharmacia.

Isolation of Restriction Fragments

All restriction endonuclease digestions were performed under conditions specified by the suppliers. Fragments of interest were separated electrophoretically on 5, 6 or 7% polyacrylamide slab gels, detected by 300 nm ultraviolet light after ethidium bromide staining, and extracted as suggested by Maxam and Gilbert (10,11).

Preparation of Plasmids

λHt 4 was isolated from a human-λ recombinant library (12) provided by T. Maniatis, using standard procedures (R. Buckland, H. Cooke, K.L. Roy, J. Dahlberg and E. Lund, in preparation). Briefly, plaques were screened by the Benton and Davis (13) method using nick-translated DNA from the plasmid pXt 210 (2) as a probe. Several positive plaques were picked, purified, and grown in quantity. DNA was isolated from the recombinant phage and characterized by restriction endonuclease digestion (Eco RI, Bam HI, Bgl II, and Hind III). Southern transfers (14) were hybridized to pXt 210 DNA, and separately to tRNA labeled, after partial thermal degradation, with $\gamma[^{32}P]$-ATP using polynucleotide kinase.

Recombinant plasmids containing segments of the human DNA insert in λHt 4 were prepared by ligation of restriction fragments into the cloning vehicle pAT 153 (15), followed by transformation of *E. coli* HB 101 (16)

with the ligated DNA. The desired transformants were detected by small scale plasmid isolation (17) after selection of $amp^r tet^s$ or $amp^r tet^r$ colonies. Restriction endonuclease digestions verified the identity of the inserted segments and, if necessary, their orientation.

End-Labeling of Restriction Fragments

All DNA fragments were labeled at their 3'-termini by using the Klenow fragment of E. coli DNA polymerase I and appropriate $\alpha[^{32}P]$ labeled and unlabeled deoxyribonucleoside triphosphates (2,19). Flush ends were labeled at 25° (to allow exonuclease removal of the ultimate nucleotide) and staggered ends (with 5' extensions) were labeled at 15°. After labeling, the DNA was precipitated by addition of ammonium acetate to 2M, followed by three volumes of ethanol, as described by Maxam and Gilbert (11). Cleavage with an appropriate restriction endonuclease was used in most cases to separate the two labeled ends, except in a few cases where strand separation on polyacrylamide slab gels was successful.

DNA Sequencing

The chemical sequencing procedures of Maxam and Gilbert (10,11) were used throughout with little modification. In some experiments the T-specific reaction described by Friedman and Brown (20) was included, and more recently the T-specific reaction of Rubin and Schmid (21) has been used. Best results were obtained when a fragment was chromatographed on DEAE-Sephacel after elution from polyacrylamide gel. Electrophoretic separation of the end-labeled fragments after completion of the sequencing reactions was achieved with thin (0.25, 0.4, 0.5 mm) polyacrylamide (6,7,8,12,16,20%) gels (22) using the 8.3 M urea-100 mM tris-borate-EDTA buffer system of Maxam & Gilbert (11). In some cases the modified electrophoretic separation technique described by Smith and Calvo (23) was used. Visualization of radioactive bands was either radioautographic at -20° using Kodak NS-54T film or, more commonly, fluorographic at -90° using Kodak XR-1 film and Dupont Cronex Lightning Plus intensifying screens (24).

RESULTS

Sequencing Strategy

Most of the human insert of λHt 4 was subcloned as a series of overlapping fragments in the plasmid pAT 153 (15). Figure 1 shows the relationship of the recombinant plasmids obtained to the insert in λHt 4.

Sequencing was conducted exclusively by the method of Maxam and Gilbert (10,11) using only 3' terminal labeling (2,19). All regions of the human

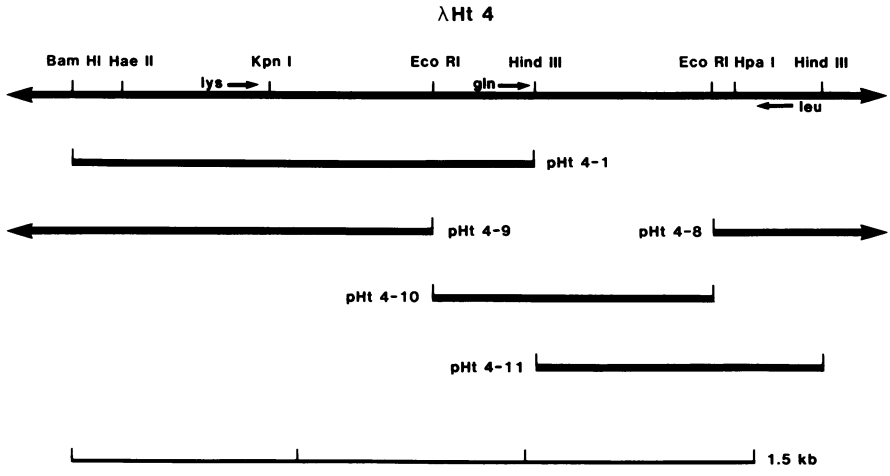

Figure 1. Restriction map of the tRNA gene-containing region of λHt 4. Sites are shown for the major six-base-specific restriction endonucleases; the segments subcloned for sequence analysis are also indicated. The locations and orientations of the three tRNA genes are shown by short arrows.

DNA were sequenced at least twice, and on both strands over most of the sequence. Restriction endonuclease digestion of the intact plasmids revealed that Alu I, Dde I, Hinf I, Msp I, Rsa I, Sau 3AI and Taq I gave products indicative of cleavage within the insert. A restriction map of the segment shown in Fig. 1, together with the major sequencing reactions, is shown in Fig. 2, and the complete sequence is shown in Fig. 3. The variety of restriction sites labeled, and the use, in a number of cases, of more than one enzyme to cut labeled fragments, gives us considerable confidence in the sequence described. With one exception, all restriction sites predicted by the sequence were cut by the appropriate enzyme, of those we had available for testing. The enzymes not tested include Bbv I, Hae III, Mbo II, Mnl I and Tha I.

The single exception was the Ava II site centered at 390. This site overlaps an Eco RII site, and it is thus assumed that the methylation of the C at position 392 also blocks Ava II cleavage. This would be analogous to the case reported by Backman (25), and also similar to the $tRNA^{tyr}$ gene described by Müller and Clarkson (2) where modification of Mbo I sites blocked cleavage at overlapping Taq I sites. Protection of restriction sites by methylation of overlapping sites seems to be fairly common (26).

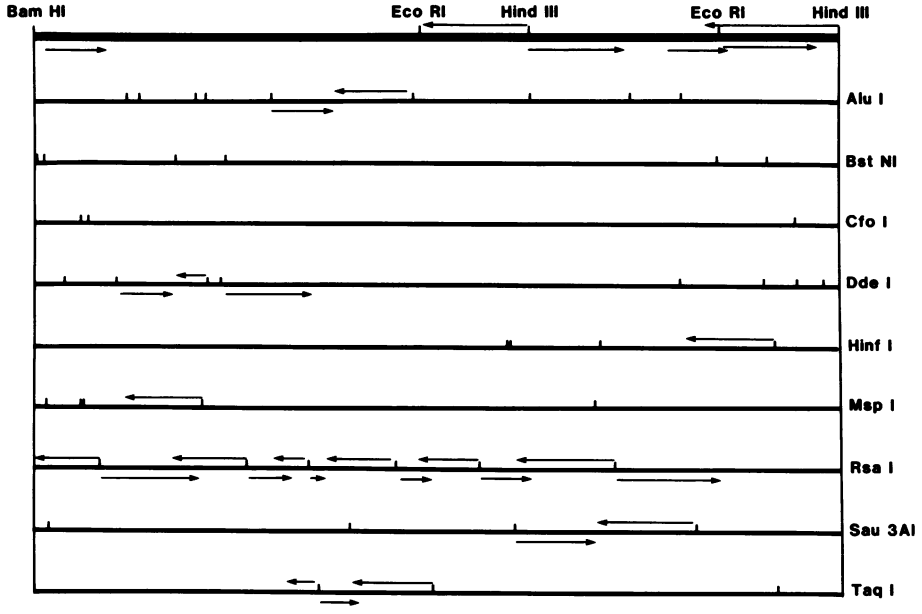

Figure 2. End-labeling and secondary cleavage restriction sites of the tRNA gene-containing region of λHt 4. The arrows indicate the direction and distance sequenced. The four and five-base-specific restriction endonuclease sites shown are consistent with their use as end-labeling sites and/or secondary cleavage sites.

tRNA Genes

A segment of cloned human DNA between nucleotides 343 and 417 appears to be a gene for a $tRNA^{lys}$. This gene would direct the synthesis of the product shown in Fig. 4a, with the CCA terminus (not encoded in the DNA) included for completeness. The $tRNA^{lys}$ shown in Fig. 4a has the same sequence as the rabbit liver $tRNA_3^{lys}$ (27) prior to secondary modifications. The $tRNA^{lys}$ gene has been shown to function efficiently *in vivo* on injection into *Xenopus oocytes* (28). Furthermore, *in vitro* site-directed mutagenesis has been used to convert this gene (transferred from pAT 153 to the bacteriophage M13 mp7) to a functional nonsense (UAG) suppressor gene (28).

Another segment of the insert appears to be a gene for a $tRNA^{gln}$ capable of recognizing the codon CAG. This presumed gene, between nucleotides 940 and 1013, also has the usual base pairing and invariant nucleotides, and would be expected to direct the synthesis of the product shown in Fig. 4b. This gene also has been shown to function upon injection into *Xenopus* oocyte nuclei (28) and can be mutated to an amber suppressor gene.

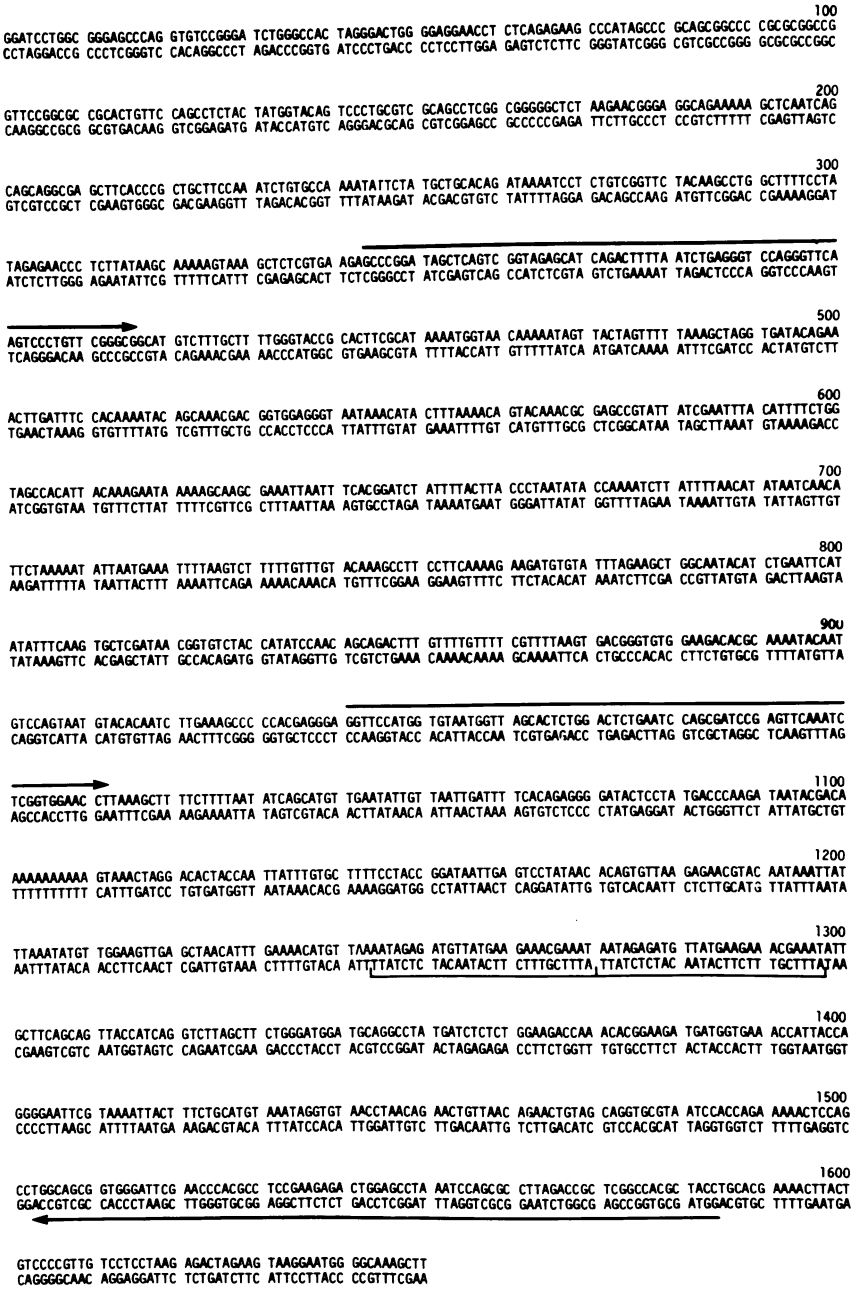

Figure 3. The sequence of the tRNA gene-containing region of λHt 4. The orientation and extent of the tRNA genes is indicated by arrows. The twenty seven base pair tandem repeat is also shown.

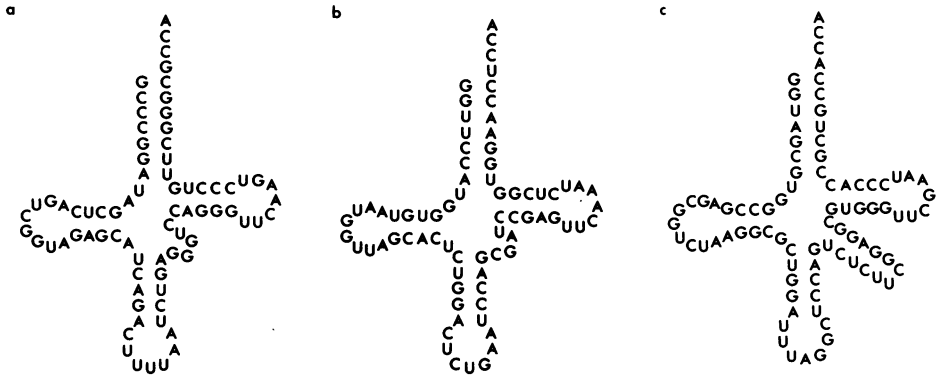

Figure 4. The sequences of the tRNAs encoded by λHt 4. (a) the $tRNA^{lys}$, (b) the $tRNA^{gln}$, (c) the $tRNA^{leu}$. In each case only the CCA terminus (not encoded by the DNA) has been added; none of the secondary modifications are indicated.

Between base pairs 1502 and 1585, but on the opposite DNA strand, a $tRNA^{leu}$ gene recognizing the codons CUA and CUG is located. Its unmodified sequence is shown in Fig. 4c, and it differs in only two positions from the bovine liver $tRNA^{leu}$ (specific for CUU, C and A) reported by Pirtle et al. (29). This tRNA gene has not yet been tested for transcriptional activity in Xenopus oocyte nuclei.

Examination of the 1.65 kb human DNA sequence using the tRNA program of Staden (30) confirmed the observation that only three tRNA genes are present.

In Vitro Transcription

Each of the tRNA genes has been tested in a HeLa cell lysate for the ability to direct synthesis of a mature sized tRNA product. As shown in Fig. 5 all three tRNA genes produce precursors which, on prolonged incubation, are converted to products with electrophoretic mobilities indicative of mature tRNAs. The products have not been analyzed for secondary modifications

Flanking Sequences

The 3' flanking sequence of the $tRNA^{lys}$ gene contains two T-rich regions, generally thought to be RNA polymerase III termination signals (31). Two similar T-rich regions are found close to the 3' end of the $tRNA^{gln}$ gene. The $tRNA^{leu}$ gene appears to have a single block of five Ts as a termination signal. All three tRNA genes have what may be secondary terminators 30 to 75 base pairs to the 3' side, in each case a run of four or five Ts (Fig. 6).

The 5' flanking sequences show only limited homology and little resem-

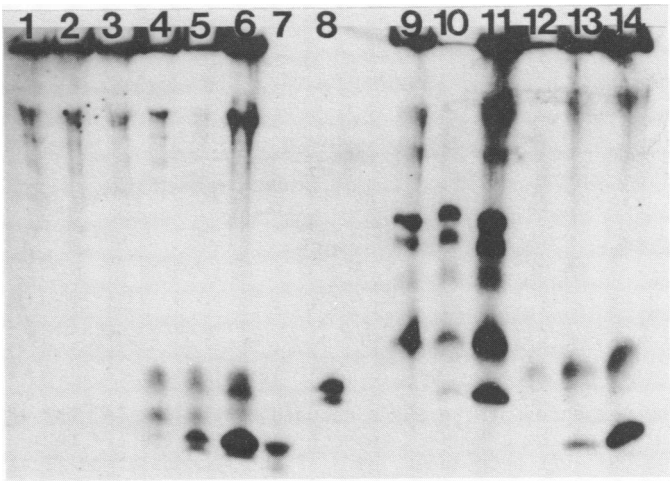

Figure 5. In vitro transcription of three human tRNA genes. The plasmids pAT 153 (lanes 1-3), pHt 4-10 ($tRNA^{Gln}$ gene, lanes 4-6), pHt 4-11 ($tRNA^{Leu}$ gene, lanes 9-11), and pHt 4-12 ($tRNA^{Lys}$ gene, Bam HI to Eco RI, lanes 12-14) were incubated with HeLa cell extract (NEN) and $\alpha[^{32}P]$-GTP for 30 min (lanes 1,4,9, 12), 1 hr (lanes 2,5,10,13), or 2 hr (lanes 3,6,11,14) at 30°. Each reaction (25 µl) was stopped by the addition of 100 µl of 0.3 M NaOAc containing 1% SDS and 5 µg of E. coli carrier tRNA. Phenol-chloroform (1:1) extraction (2x) was followed by alcohol precipitation, gel electrophoresis (11.4% acrylamide and 0.6% bisacrylamide) as in ref 11, and radioautography. Lane 7 is a marker of ^{32}pCp-end labeled tRNA ($tRNA^{Gln}$) isolated by hybridization to pHt 4-10 immobilized on nitrocellulose (33). Lane 8 contains similar tRNA isolated by hybridization to immobilized pHt 4-11.

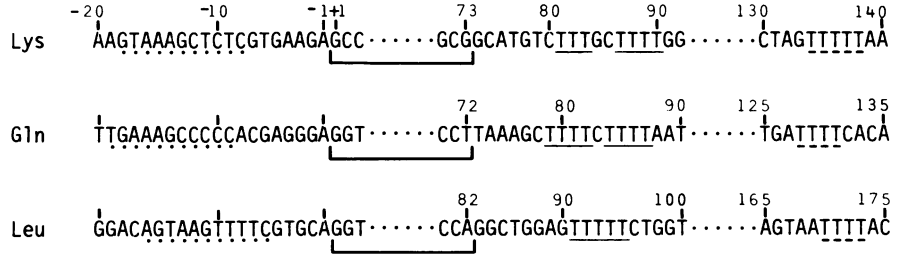

Figure 6. The flanking regions of the tRNA genes. The non-coding strands of the three tRNA genes are aligned at their 5' ends to demonstrate the limited homology of the 5' flanking sequences (dotted underlining). The termination signals are underlined, and potential secondary termination signals are also shown (dashed underlining). The segments corresponding to the mature tRNA sequences are bracketed.

blance to the 5' flanking sequences of most other tRNA genes from higher eucaryotes. All three of these human tRNA genes have a block of five pyrimidines flanked on the 5' side by a purine-rich region. This is illustrated in Fig. 6, primarily to emphasize a tentative relationship to the regulatory sequence described by DeFranco et al. (32).

The DNA between the tRNA genes does not appear to contain information sufficient to code for a protein in any of the reading frames. This is probably true also of the region to the left of the $tRNA^{lys}$ gene, but one reading frame is open from the Bam HI site to a termination codon at position 262. A feature worth noting is a perfect tandem repeat of 27 base pairs between nucleotides 1243 and 1298, approximately midway between the $tRNA^{gln}$ gene and the $tRNA^{leu}$ gene. The significance, if any, of this repeat is unknown.

DISCUSSION

The tRNA gene region of a human-λ recombinant has been sequenced and shown to contain a cluster of three tRNA genes, each of which is active in an in vitro transcription system. As in the study reported by Sekiya et al. (7) the genes are separated by about .4-.5 kb of non-coding DNA. This human gene cluster is not tandemly repeated, as is the 3.18 kb Xenopus tRNA gene cluster studied by Müller and Clarkson (2), since the original human-λ recombinant, λHt 4, has no other tRNA genes in 11 kb of human DNA (8 kb on one side, and 3 kb on the other side of the sequenced region). Endonuclease digests of human DNA have not yet been examined to determine if this cluster is repeated as a dispersed unit in the human genome.

ACKNOWLEDGEMENTS

We thank Ms. Debra McIntyre for skilled technical assistance, Dr. D. Finnegan for the 3' end labeling procedure, and Drs. E. Lund, J. Dahlberg and E. Southern for advice and suggestions in the early stages of this work. K.L.R. wishes to thank Dr. E. Southern and Prof. P.M.B. Walker for their hospitality in providing the opportunity to work at the Mammalian Genome Unit. Research conducted at the University of Alberta was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. Sprinzl, M. and Gauss, D.H. (1982a) Nucleic Acids Res. 10, r57-r81.
2. Müller, F. and Clarkson, S.G. (1980) Cell 19, 345-353.
3. Telford, J.L., Kressmann, A., Koski, R.A., Grosschedl, R., Müller, F., Clarkson, S.G. and Birnstiel, M.L. (1979) Proc. Nat. Acad. Sci. U.S. 76, 2590-2594.
4. Hofstetter, H., Kressmann, A. and Birnstiel, M.L. (1981) Cell 24, 573-585.
5. Galli, G., Hofstetter, H. and Birnstiel, M.L. (1981) Nature 294, 626-631.
6. Santos, T. and Zasloff, M. (1981) Cell 23, 699-709.
7. Sekiya, T., Kuchino, Y. and Nishimura, S. (1981) Nucleic Acids Res. 9, 2239-2250.
8. Sprinzl, M. and Gauss, D.H. (1982b) Nucleic Acids Res. 10, r1-r55.
9. Lasser-Weiss, M., Bawnik, N., Rosen, A., Sarid, S. and Daniel, V. (1981) Nucleic Acids Res. 9, 5965-5978.
10. Maxam, A. and Gilbert, W. (1977) Proc. Nat. Acad. Sci. U.S. 74, 560-564.
11. Maxam, A.M. and Gilbert, W. (1980) in Methods in Enzymology, Grossman, L. and Moldave, K., eds., Vol. 65, pp. 499-560, Academic Press, New York.
12. Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G. and Maniatis, T. (1978) Cell 15, 1157-1174.
13. Benton, W.D. and Davis, R.W. (1977) Science 196, 180-182.
14. Southern, E. (1975) J. Mol. Biol. 98, 503-517.
15. Twigg, A.J. and Sherratt, D. (1980) Nature 283, 216-218.
16. Bolivar, F. and Backman, K. (1979) in Methods in Enzymology, Wu, R., ed., Vol. 68, pp. 245-267, Academic Press, New York.
17. Birnboim, H.C. and Doly, J. (1979) Nucleic Acids Res. 7, 1513-1523.
18. Clewell, D.B. and Helinski, D.R. (1969) Proc. Nat. Acad. Sci. U.S. 62, 1159-1166.
19. Smith, M., Leung, D., Gillam, S., Astell, C., Montgomery, L. and Hall, B.D. (1979) Cell 16, 753-761.
20. Friedmann, T. and Brown, D.M. (1978) Nucleic Acids Res. 5, 615-622.
21. Rubin, C.M. and Schmid, C.W. (1980) Nucleic Acids Res. 8, 4613-4619.
22. Sanger, F. and Coulson, A.R. (1978) FEBS Lett. 87, 107-110.
23. Smith, D.R. and Calvo, J.M. (1980) Nucleic Acids Res. 8, 2255-2274.
24. Laskey, R.A. and Mills, A.D. (1977) FEBS Lett. 82, 314-316.
25. Backman, K. (1980) Gene 11, 169-171.
26. McClelland, M. (1981) Nucleic Acids Res. 9, 5859-5866.
27. Raba, M., Limburg, K., Burghagen, M., Katze, J.R., Simsek, M., Heckman, J.E., RajBhandary, U.L. and Gross, H.J. (1979) Eur. J. Biochem. 97, 305-318.
28. Temple, G., Dozy, A., Roy, K.L. and Kan, Y.W. (1982) Nature 296, 537-540.
29. Pirtle, R., Kashdan, M., Pirtle, I. and Dudock, B. (1980) Nucleic Acids Res. 8, 805-815.
30. Staden, R. (1980) Nucleic Acids Res. 8, 817-825.
31. Bogenhagen, D.F. and Brown, D.D. (1981) Cell 24, 261-270.
32. DeFranco, D., Sharp, S. and Söll, D. (1981) J. Biol. Chem. 256, 12424-12429.
33. Weiner, A. (1980) Cell 22, 209-218.