Volume 232, number 1, 187–192 FEB 05851 May 1988

Nucleotide sequence of the androgen-dependent arginine esterase mRNA of canine prostate

Pierre Chapdelaine, My-Anh Ho-Kim, Roland R. Tremblay and Jean Y. Dubé

Laboratory of Hormonal Bioregulation, Laval University Hospital Centre, 2705 Laurier Boulevard, Sainte-Foy, Québec G1V 4G2, Canada

Received 25 February 1988

The nucleotide sequence of canine prostate arginine esterase mRNA was determined using a 400 bp cDNA clone and primer-extended cDNA transcripts for the 5'-coding and noncoding regions. The mRNA contains 864 nucleotides encoding a protein of 236 amino acids preceded by 24 amino acids which constitutes both the signal and the zymogen peptides. The sequence indicates the presence of one potential glycosylation site. A high degree of homology was found between the canine enzyme and other members of the kallikrein family including human prostate specific antigen. The protein appears to be specified by a single gene.

Kallikrein; cDNA clone; RNA sequencing; Androgen-dependent protein

1. INTRODUCTION

Male reproductive tissues contain a variety of proteolytic enzymes [1]. This is particularly evident in the dog in which more than 90% of the seminal plasma proteins are constituted by a proteolytic enzyme, arginine esterase, derived from the prostate [2,3]. The human prostate and seminal plasma also contain a relatively abundant protease, prostate-specific antigen (PSA) [4]. This protein is better characterized than the canine enzyme. Its primary structure has recently been published independently by 2 groups [5,6] who showed that it has extensive homology with several proteases of the kallikrein family. One of the presumed functions of this enzyme is to hydrolyze the high-molecular-mass seminal vesicle protein thus leading to the liquefaction of the seminal coagulum [7]. By contrast, there is no seminal vesicle and no formation of a seminal coagulum in the canine species. The physiological substrate of arginine esterase is thus at present unknown, since the enzyme is unable to generate vasoactive kinins from kininogen [8] like the other members of the kallikrein family. This protein is highly dependent on androgenic hormones [9] and constitutes a valuable model system for the study of hormone action at the molecular level. A partial amino acid sequence of both the light and heavy chains of arginine esterase has been published by our laboratory [10] and we have also recently described a 400 bp cDNA clone coding for this protein [11]. Here, we propose the complete nucleotide sequence of its mRNA and the deduced amino acid sequence.

Correspondence address: J.Y. Dubé, Laboratory of Hormonal Bioregulation, Laval University Hospital Centre, 2705 Laurier Boulevard, Sainte-Foy, Québec G1V 4G2, Canada

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession no. Y00751

2. MATERIALS AND METHODS

2.1. *Recombinant clone*

Plasmid pAgE400 containing an incomplete cDNA coding for canine prostate arginine esterase was isolated and sequenced as in [11].

2.2. *Oligonucleotide primers*

In order to determine the sequence of the 5'-portion of the

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

mRNA not encoded in the cDNA clone pAgE400, we used the RNA sequencing and cDNA sequencing procedures of Geliebter et al. [12] based on the utilization of complementary oligonucleotide primers and primer extension. The whole sequence strategy is illustrated in fig.1. The oligonucleotide primers were chemically synthesized using a Biosearch model 8700 DNA synthesizer (New Brunswick Scientific). These primers were labeled with [γ-^{32}P]ATP (Amersham, 5000 Ci/mmol) using T_4 polynucleotide kinase (Bethesda Research Laboratories) as described by Geliebter et al. [12]. The first primer 5′-CCCTGGGTGAAAAATCG-3′ was complementary to the mRNA sequence from nucleotides 523 to 539 in fig.2. With the first primer extension reaction, an additional 200 nucleotides were read and thus a second oligonucleotide primer could be devised and so on. Table 1 lists all the primers that were synthesized for sequencing of both RNA and cDNA.

2.3. *RNA sequencing*

5 ng ^{32}P-labeled oligonucleotide primer and 2 μg poly(A)$^+$ RNA prepared as described by Chirgwin et al. [13] were heated for 3 min at 80°C in 10 μl annealing buffer (250 mM KCl/10 mM Tris-HCl buffer, pH 8.3) and then for 45 min at a different temperature for each oligonucleotide according to the formula: T (°C) = $4(G+C) + 2(A+T) - 5$. The mixture was then left standing at room temperature for 30 min. The reverse transcription reaction was performed as described [12] with the exception that 7-deaza-2′-deoxyguanosine 5′-triphosphate from Pharmacia was substituted for dGTP. Furthermore, RAV-2 reverse transcriptase from Amersham was used with final concentrations of each ddNTP from Pharmacia (0.16 mM ddATP; 0.08 mM ddGTP, ddCTP and ddTTP).

2.4. *cDNA sequencing*

The annealing reaction was performed with 5 μg canine prostate poly(A)$^+$ RNA in the presence of 25–50 ng of one oligonucleotide primer (1, 2, 3 or 4) in a volume of 100 μl. The reverse transcription reaction was then performed at 50°C for 45 min with 150 μl reverse transcription buffer [12] containing 35–40 U RAV-2 reverse transcriptase. The cDNA generated by primer extension was extracted sequentially with equal volumes of phenol, then phenol/chloroform (1:1) and finally chloroform. The DNA dissolved in 0.4 M ammonium acetate was precipitated 3 times with ethanol. The RNA:cDNA hybrids were denatured in boiling water for 5 min. Thereafter ^{32}P-labeled oligonucleotide primers (5, 6, 7 or 8) were annealed to the cDNA and were extended with DNA polymerase I (Klenow fragment) from Amersham in the presence of dNTPs and

Table 1

Oligonucleotide primers for mRNA and cDNA sequencing

Oligonucleotide primers (5′ ⟶ 3′)	Corresponding position in mRNA sequence	Annealing temperature (°C)
1 CCCTGGGTGAAAAATCG	539–523	47
2 TCTATCTTCCCCTGGACGGAT	383–363	59
3 CCTCACAATTACTGTTTGCAC	252–232	55
4 GCCTCCTATGATCCGGGGCTG	118–98	65
5 CTCACTGTCTGCCAGCT	1–17	49
6 TCTCAGCCCTGGCAGGTGGCT	138–158	55
7 GAAGATGAAGGCCAGTTAGTC	285–305	57
8 GCTGTGAGGGTGATGGACCTG	435–455	63

Primers 1–4 were complementary to the mRNA sequence and were used for mRNA sequencing while primers 5–8 were complementary to the cDNA sequence and were used for cDNA sequencing

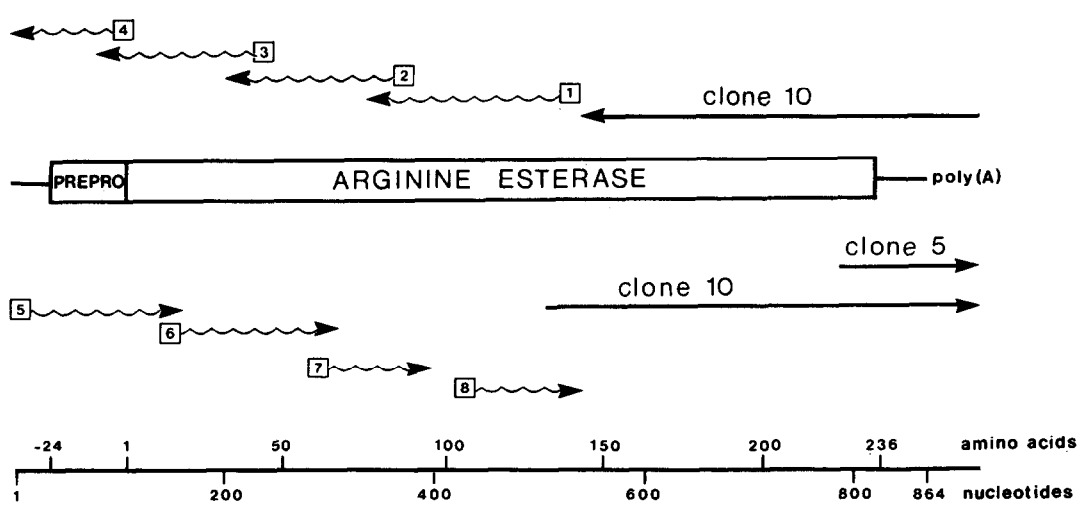

Fig.1. Sequencing strategy for canine prostate arginine esterase mRNA. The horizontal rectangle shows the amino acid coding regions including the signal peptide (PRE) and putative activation peptide (PRO) while the lines extending from the rectangle represent the lengths of the 5′- and 3′-noncoding regions. The horizontal arrows indicate the direction and extension of each sequencing run with clones 10 and 5 or by primer extension (wavy lines) with primers 1–4 (mRNA sequencing) and primers 5–8 (cDNA sequencing).

ddNTPs [12]. The gels used for RNA and cDNA sequencing contained 5% acrylamide and 7 M urea. However, in some cases in which more details were needed in the vicinity of the primer, 10 and 15% polyacrylamide gels were also used.

2.5. *Southern blot analysis*

Blood DNA was prepared as described by Gustafson et al. [14]. It was then completely digested with *Eco*RI, *Bam*HI, *Pst*I, *Hin*dIII, *Hae*III and *Kpn*I. Samples containing 10 μg DNA were electrophoresed in gels of 1% agarose [15] and transferred to Hybond-N membranes as suggested by the supplier (Amersham). After the transfer, the filters were prehybridized in 50% formamide hybridization buffer for 5 h at 42°C. The insert of the prostatic cDNA clone pAgE400 [11] was labeled with the multiprime labeling system of Amersham and hybridization was carried out at 42°C for 36 h in 6 × STE, 1 × Denhart's solution, 50% formamide, 0.1% SDS and 100 μg/ml of salmon sperm DNA. The filters were washed 3 times with 1 × STE, 0.2% SDS at room temperature and then twice for 15 min at 68°C with 1 × STE, 0.2% SDS. The filters were air-dried and exposed to Kodak X-AR2 X-ray films at −80°C.

3. RESULTS

3.1. *Nucleotide sequence of arginine esterase mRNA*

The combination of sequencing of clone pAgE400 and of RNA and uncloned cDNA provided all the information for obtaining the complete nucleotide sequence of arginine esterase mRNA (fig.2). The entire length of the mRNA excluding the poly(A) tail comprises 864 nucleotides encoding a protein of 236 amino acids (M_r 26003) preceded by 24 amino acids constituting both the signal and zymogen peptides. The sequence also includes 35 nucleotides in the 5′-noncoding region and 46 nucleotides in the 3′-noncoding region after the UGA stop codon. The deduced amino acid sequence shows the presence of the catalytic

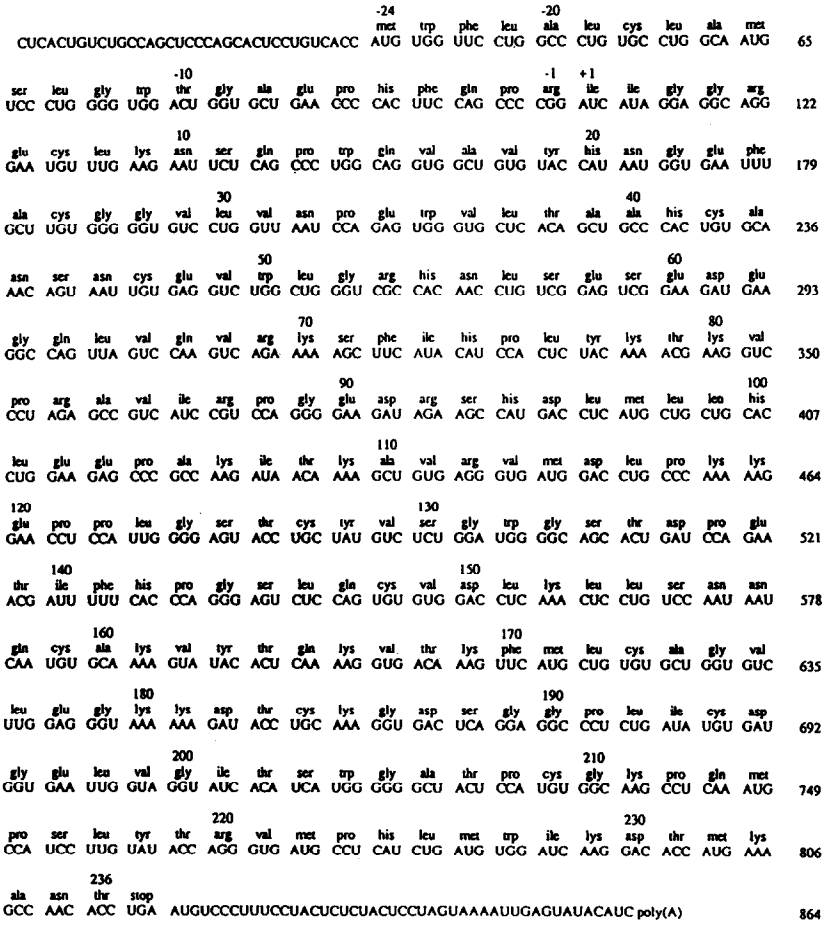

Fig.2. Complete nucleotide sequence of canine prostatic arginine esterase mRNA and deduced amino acid sequence of the encoded preproenzyme. Amino acid numbering starts with the first residue of the mature enzyme [10].

triad (His 41, Asp 95, Ser 188) of serine protease and confirms that arginine esterase belongs to that class of enzymes. The amino acid sequence presents a high degree of homology with other kallikreins and in particular with human prostate specific antigen (58%), γ-subunit of mouse nerve growth factor (56%) and human pancreatic kallikrein (55%).

One particular feature of the nucleotide sequence is the presence, 15 nucleotides upstream from the poly(A) stretch, of a putative polyadenylation signal consisting of AGUAAA as in human PSA [6], human pancreas kallikrein [16] and human prostatic secretory protein of 94 amino acids [17] but different from the more canonical AAUAAA. The amino acid sequence Asn-Leu-Ser at positions 55–57 indicates the possibility of one *N*-linked glycosylation site.

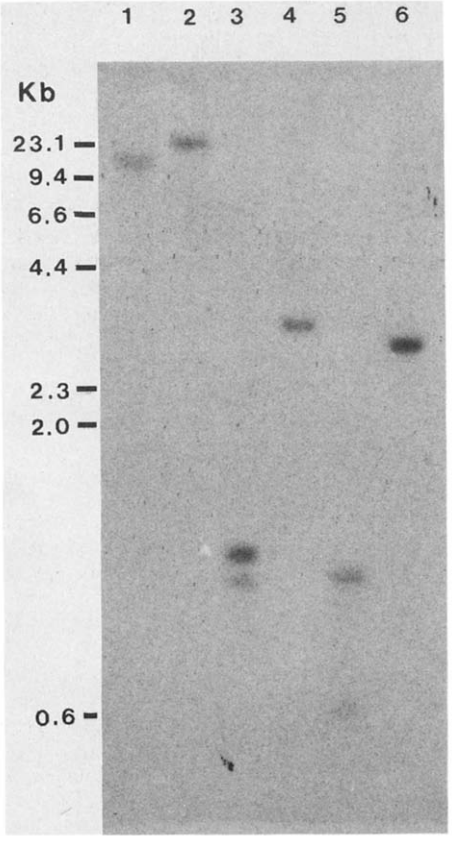

Fig.3. Southern blot analysis of arginine esterase gene. Canine blood DNA was digested with *Eco*RI (1), *Bam*HI (2), *Pst*I (3), *Hin*dIII (4), *Hae*III (5) and *Kpn*I (6) and run on a 1% agarose gel. The DNA was transferred to Hybond-N membrane and hybridized with ^{32}P-labeled pAgE400 cDNA insert as described in section 2.

The amino acid sequence at positions 1–26 and 84–116 was found to be identical to the partial amino acid sequence of the NH_2-terminal portion of the heavy and light chains of native seminal plasma arginine esterase [10].

3.2. *Southern blot analysis*

In order to evaluate the number of arginine esterase-related genes in the dog genome, high-molecular-mass DNA from blood was digested with restriction enzymes and was analyzed by Southern blot analysis using the 400 bp cDNA clone as a probe (fig.3). A single strong hybridization band was observed after digestion of the DNA with *Hin*dIII, *Eco*RI, *Bam*HI or *Kpn*I while 2 bands were present after *Pst*I digestion and *Hae*III.

4. DISCUSSION

The present results are in complete agreement with our previous work on the partial amino acid sequence of arginine esterase from canine seminal plasma [10]. Furthermore, they extend those we obtained with the partial cDNA clone pAgE400 [11]. They also illustrate the fact that the sequencing of an abundant mRNA can be carried out accurately without cloning once a small portion of the nucleotide sequence is known or can be deduced. An additional reason for the success of this approach is probably due to the absence of other closely related mRNAs in canine prostate.

The total length of arginine esterase mRNA is quite similar to those of rat kallikrein mRNA [18], human pancreatic kallikrein mRNA [16], human cathepsin G [19] and rat pancreatic trypsinogens [20]. In contrast, human PSA mRNA is significantly longer (1.5 kb), particularly in its 3′-untranslated region [6]. In view of the high levels of conservation (90%) of the 3′-untranslated region in the various rat kallikrein mRNAs, it has been suggested that this region could have important functions [18]. However, these observations cannot presently be interpreted as an indication that the 3′-untranslated regions of human PSA and canine arginine esterase mRNAs have different functions.

The amino acid sequence shows many

similarities and also some differences between canine arginine esterase and other members of the kallikrein family. The prepropeptide sequence contains 24 amino acids like most of the other kallikreins. The boundary between the signal peptide and the zymogen peptide cannot be predicted with certainty. However, according to Von Heijne's criteria [21], the probable cleavage site of the signal peptidase should be after the alanine at −8. An additional possible prepeptide cleavage site has also been proposed after the glycine at −12 for several kallikreins [18] on the basis of the structural criteria elaborated by Perlman and Halvorson [22]. The amino acid sequence of the arginine esterase prepropeptide is much less conserved than those of other kallikreins [18]. Another uncertainty is the exact extent of the autolysis loop. From the amino acid analysis previously reported [10], native arginine esterase heavy chain begins with the alanine at position 84. By taking into consideration the amino acid composition of the native two-chain enzyme [10] and this sequence, it can be calculated that the autolysis loop should involve 11–13 amino acids.

The potential glycosylation site at position 55 is in accord with previous results showing that only the light chain of arginine esterase was glycosylated [2,3]. This site is located at a different position from those observed in mouse, rat and human kallikreins. It is also different from the glycosylation site of human PSA at position 47 [6].

The amino acid sequence also provides useful information on substrate specificity. The presence of an aspartate residue at position 182 indicates the trypsin-like specificity of the enzyme [23]. This result is in accord with the observed substrate specificity with synthetic substrates [8]. By contrast, PSA has a serine at the equivalent position which is indicative of chymotrypsin-like specificity [6]. Also worthy of note is that Gly 205 and Ser 216 are present at homologous positions in canine prostatic arginine esterase as well as in human pancreatic kallikrein [16] and human PSA [6]. These amino acids are believed to allow access to the catalytic pocket of amino acid residues with bulky side chains [23]. Two other amino acids, Tyr 92 and Trp 204, are presumed to be important determinants for the hydrolysis of plasma kininogen and release of vasoactive kinins [24]. In arginine esterase, Trp 204 is present but Tyr 92 has been replaced by an arginine residue. These results could thus explain the observed absence of kininogenase activity of canine arginine esterase [8,25].

The results obtained with the Southern blot analysis strongly suggest that a single gene of arginine esterase is present in the canine genome and that it is not part of a large kallikrein gene family such as that found in mouse [26] and rat [27]. In the genome of the mouse, the 24 kallikrein-related genes are highly homologous and are clustered on chromosome 7 [28]. In man, the kallikrein family may be smaller with a maximum of 3 members [29], respectively. The absence of genes highly homologous to canine prostate arginine esterase in Southern blot hybridization experiments does not exclude the possibility of a larger kallikrein family in that species or that all these genes could be clustered on a single gene locus. These questions will await the cloning of arginine esterase gene and of other kallikrein-related genes in the dog.

Acknowledgements: This work was supported by a grant from the Medical Research Council of Canada (no. MT-4780). The authors wish to thank Mrs Lucie Turcotte for secretarial assistance and in particular for the preparation of fig.2. They are also grateful to Dr Guy Bellemare and Mr Claude Potvin, Biochemistry Department, Laval University, for helpful suggestions.

REFERENCES

[1] Zaneveld, L.J.D., Polakoski, K.L. and Schumacher, G.F.B. (1975) in: Proteases and Biological Control (Reich, E.D. et al. eds) pp.683–706, Cold Spring Harbor Laboratory, NY.
[2] Isaacs, W.B. and Shaper, J.H. (1983) J. Biol. Chem. 258, 6610–6615.
[3] Chapdelaine, P., Dubé, J.Y., Frenette, G. and Tremblay, R.R. (1984) J. Androl. 5, 206–210.
[4] Ban, Y., Wang, M.C., Loor, R. and Chu, M. (1984) Biochem. Biophys. Res. Commun. 123, 482–488.
[5] Watt, K.W.K., Lee, P.J., Mtimkulu, T., Chan, W.P. and Loor, R. (1986) Proc. Natl. Acad. Sci. USA 83, 3166–3170.
[6] Lundwall, A. and Lilja, H. (1987) FEBS Lett. 214, 317–322.
[7] Lilja, H. (1985) J. Clin. Invest. 76, 1899–1903.
[8] Frenette, G., Dubé, J.Y. and Tremblay, R.R. (1985) Biochim. Biophys. Acta 838, 270–276.
[9] Dubé, J.Y., Chapdelaine, P. and Tremblay, R.R. (1983) Can. J. Biochem. Cell Biol. 61, 756–763.

[10] Lazure, C., Leduc, R., Seidah, N.G., Chrétien, M., Dubé, J.Y., Chapdelaine, P., Frenette, G., Paquin, R. and Tremblay, R.R. (1984) FEBS Lett. 175, 1–7.
[11] Chapdelaine, P., Potvin, C., Ho-Kim, M.A., Larouche, L., Bellemare, G., Tremblay, R.R. and Dubé, J.Y. (1988) Mol. Cell. Endocrinol., in press.
[12] Geliebter, J., Zeff, R.A., Melvold, R.W. and Nathenson, S.G. (1986) Proc. Natl. Acad. Sci. USA 83, 3371–3375.
[13] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[14] Gustafson, S., Proper, J.A., Bowie, E.J.W. and Sommer, S.S. (1987) Anal. Biochem. 165, 294–299.
[15] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in: Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[16] Fukushima, D., Kitamura, N. and Nakanishi, S. (1985) Biochemistry 24, 8037–8043.
[17] Mbikay, M., Nolet, S., Fournier, S., Benjannet, S., Chapdelaine, P., Paradis, G., Dubé, J.Y., Tremblay, R.R., Lazure, C., Seidah, N.G. and Chrétien, M. (1987) DNA 6, 23–29.
[18] Ashley, P. and MacDonald, R.J. (1985) Biochemistry 24, 4512–4520.
[19] Salvesen, G., Farley, D., Shuman, J., Przybyla, A., Reilly, C. and Travis, J. (1987) Biochemistry 26, 2289–2293.
[20] MacDonald, R.J., Stary, S.J. and Swift, G.H. (1982) J. Biol. Chem. 257, 9724–9732.
[21] Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4690.
[22] Perlman, D. and Halvorson, H.O. (1983) J. Mol. Biol. 167, 391–409.
[23] Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt-Kastner, G. and Bartunik, H. (1983) J. Mol. Biol. 164, 237–282.
[24] Chen, Z. and Bode, W. (1983) J. Mol. Biol. 164, 283–311.
[25] Bhoola, K.D., Morley, J., Shacter, M. and Smaje, L.H. (1965) J. Physiol. 179, 172–184.
[26] Evans, B.A., Drinkwater, C.C. and Richards, R.I. (1987) J. Biol. Chem. 262, 8027–8034.
[27] Gerald, W.L., Chao, J. and Chao, L. (1986) Biochim. Biophys. Acta 866, 1–14.
[28] Mason, A.J., Evans, B.A., Cox, D.R., Shine, J. and Richards, R.I. (1983) Nature 303, 300–307.
[29] Schedlich, L.J., Bennetts, B.H. and Morris, B.J. (1987) DNA 6, 429–437.