Volume 16 Number 23 1988 Nucleic Acids Research

Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells

Alec J.Jeffreys*, Victoria Wilson, Rita Neumann and John Keyte[1]

Departments of Genetics and [1]Biochemistry, University of Leicester, Leicester, LE1 7RH, UK

Received October 14, 1988; Accepted November 4, 1988

ABSTRACT

Hypervariable minisatellites can be amplified from human DNA by the polymerase chain reaction, using primers from DNA flanking the minisatellite to amplify the entire block of tandem repeat units. Minisatellite alleles up to 5-10 kb long can be faithfully amplified. At least six minisatellite loci can be co-amplified from the same DNA sample and simultaneously detected to provide a reproducible and highly variable DNA fingerprint which can be obtained from nanogram quantities of human DNA. The polymerase chain reaction can also be used to analyse single target minisatellite molecules and single human cells, despite the appearance of spurious PCR products from some hypervariable loci. DNA fingerprinting at the level of one or a few cells therefore appears possible.

INTRODUCTION

Tandem-repetitive minisatellite regions in vertebrate DNA frequently show high levels of allelic variability in the number of repeat units [1-4]. Hybridization probes capable of detecting multiple minisatellites and producing individual-specific DNA fingerprints have been developed [5-7], as well as cloned human minisatellites which provide locus-specific probes for individual hypervariable loci [5,8-10]. These highly informative genetic markers have found widespread application in many areas of genetics, including linkage analysis [9,11-13], determination of kinship in for example paternity and immigration disputes [6,10,14,15], monitoring bone marrow transplants [16,17], and for individual identification in forensic medicine [10,18,19]. Applications to typing forensic samples such as blood and semen stains or hair roots are however limited by the sensitivity of the hybridization probes, which require at least 50ng of relatively undegraded human DNA for typing with locus-specific minisatellite probes [10] and 0.1-1μg DNA for analysis with multi-locus DNA fingerprint probes [6].

The enzymatic amplification of DNA by the polymerase chain reaction (PCR) [20] enables much smaller amounts of human DNA to be analysed. The remarkable specificity of thermostable Taq polymerase has greatly simplified

PCR [21] and has allowed typing of some classes of human DNA polymorphism to be extended to single hair roots [22] and indeed to individual somatic cells and sperm [23]. In most work to date, PCR has been used to amplify short regions of human DNA, usually a few hundred basepairs long [21-23]. Base substitutional polymorphisms can be detected by hybridizing PCR products with allele-specific oligonucleotides [22,23], by DNA sequence analysis of PCR products [24], or, if the base substitution affects a restriction site, by cleavage of PCR products with a restriction endonuclease [25]. Deletion/insertion polymorphisms can likewise be analysed by sizing PCR products by gel electrophoresis [22]. Most of these marker systems are however dimorphic and their utility in for example forensic medicine is limited by their relatively low variability in human populations.

Since Taq polymerase is capable of efficiently amplifying regions of DNA up to 2 kb long [21], it therefore appeared possible that PCR could be extended to the amplification of entire human minisatellites. We now show that high fidelity amplification of hypervariable loci is possible, even from minute amounts of human DNA.

MATERIALS AND METHODS

Preparation of genomic DNA, oligonucleotides and hybridization probes

Human DNA samples were provided by CEPH, Paris, or were prepared from venous blood as described elsewhere [26]. Oligonucleotides synthesised on an Applied Biosystems 380B DNA synthesiser using reagents supplied by Cruachem were purified by ethanol precipitation and dissolved in water. The 5.6 kb Sau3A insert from human minisatellite clone λMS32 [10] was subcloned into the BamHI site of pUC13 [27]. Similarly, the minisatellite inserts from recombinant M13 RF DNAs 33.1, 33.4 and 33.6 [5], isolated as a 1.9 kb BamHI-EcoRI fragment, a 2.7 kb Sau3A-EcoRI fragment and a 0.7 kb BamHI-EcoRI fragment respectively, were subcloned into pUC13 digested with BamHI plus EcoRI to produce the plasmid series p33.1, p33.4 and p33.6. Appropriate minisatellite-containing DNA fragments were isolated from restriction endonuclease digested plasmid DNAs by electrophoresis through 1% low gelling temperature agarose (SeaPlaque); gel slices containing DNA fragments were dissolved in water at 65° to a final concentration of 2 μg/ml DNA. 10ng aliquots of DNA were labelled with ^{32}P by random oligonucleotide priming [28].

Polymerase chain reaction

Aliquots of human DNA, diluted if necessary with 5mM Tris-HCl (pH7.5) in

the presence of 0.1μM PCR oligonucleotide primers as carrier, were amplified in 10μl 67mM Tris-HCl (pH8.8), 16mM $(NH_4)_2SO_4$, 6.7mM $MgCl_2$, 10mM 2-mercaptoethanol, 6.7μM EDTA, 1.5mM dATP, 1.5mM dCTP, 1.5mM dGTP, 1.5mM dTTP (Pharmacia), 170μg/ml bovine serum albumin (DNase free, Pharmacia) plus 1μM of each oligonucleotide primer and 1.5 units Taq polymerase (Anglian Biotech). Reaction mixes in 1.5 ml microcentrifuge tubes were overlaid with 40μl paraffin oil and cycled for 1 min at 95°, 1 min at 60° and 15 min at 70° on an Intelligent Heating Block (Cambio, Cambridge). Final amplification reactions were generally chased by a final step of 1 min at 60°, to anneal any remaining single-stranded DNA with primer, followed by an extension phase at 70° for 15 min.

Southern blot analysis of PCR products

Paraffin oil was removed from PCR reactions by extraction with diethyl ether. Agarose gel electrophoresis of PCR products, Southern blotting onto Hybond-N (Amersham) and hybridization with ^{32}P-labelled minisatellite probes were carried out as decribed previously [10], except that competitor human DNA was omitted from all hybridizations. Restriction digests and S1 nuclease digestion of PCR products were performed by diluting 5μl PCR reaction with 25 μl restriction endonuclease or S1 nuclease buffer [29] and digesting for 30 min at 37° with 3 units restriction endonuclease or S1 nuclease (BRL) prior to gel electrophoresis.

Isolation and PCR analysis of single human cells

Lymphocytes were isolated by diluting venous blood with an equal volume of 1xSSC (saline sodium citrate, 0.15M NaCl, 15mM trisodium citrate, pH 7.0), layering over Histopaque-1119 (Sigma) and centrifuging at 2000g for 10 min. Cells at the interface were diluted with 3 vol 1xSSC and banded again over Histopaque. Cells were pelleted by centrifuging at 2000g for 10 min, washed three times with 1xSSC, with centrifugation, and resuspended in 1xSSC to 10^4 cells/ml.

Buccal cells were isolated by diluting 0.5ml saliva with 5ml 1xSSC and centrifuging at 2000g for 10 min. The cell pellet was rinsed three times with 1xSSC and resuspended to 10^4 cells/ml.

Approximately 0.1μl aliquots of the cell suspensions were pipetted onto a siliconised microscope slide and rapidly examined at 100x magnification on an inverted microscope. Droplets containing a single nucleated cell were immediately diluted with 0.4μl 1xSSC and transferred to an Eppendorf tube using a disposable tip pipette. The microscope slide was re-examined to check that the cell had been removed with the droplet.

Cells were lysed prior to PCR either by heating or by treatment with sodium dodecyl sulphate (SDS) and proteinase K [23]. In the former case, the cell droplet was diluted with 4.5µl 5mM Tris-HCl (pH 7.5) containing 0.1µM oligonucleotide primers, overlaid with paraffin oil and heated at 95° for 3 min prior to the addition of 5µl 2x concentrated PCR buffer/primers/Taq polymerase and amplification. In the latter case, the cell droplet was mixed with 0.5µl 5mM Tris-HCl (pH7.5), 0.1µM primers plus 1µl 5mM Tris-HCl (pH7.5), 40mM dithiothreitol, 3.4 µM SDS, 50µg/ml proteinase K [23], overlaid with paraffin oil and digested at 37° for 45 min. 3µl water were added to the digest, and heated at 95° for 3 min to inactivate proteinase K prior to addition of 5µl 2x PCR reaction mix as above.

RESULTS

Selection of human minisatellites for amplification by PCR

The strategy for amplifying minisatellites is shown in Fig.1. Oligonucleotide primers corresponding to unique sequence DNA flanking the minisatellite are used to drive amplification of the entire minisatellite by Taq polymerase. Amplified alleles are detected by Southern blot hybridization with a minisatellite probe located internal to the priming sites. Six cloned minisatellites were chosen for study (Table 1). Two of them, pλg3 and λMS32 [8,10], detect highly variable loci with heterozygosities of 97% and more than 40 alleles varying in the number of repeat units. The other four minisatellites, 33.1, 33.4 and 33.6 [5] and pMS51, isolated as a Sau3A-EcoRI DNA fragment cloned from a DNA fingerprint (A.J.Jeffreys, unpublished data), detect much less variable loci with heterozygosities of 66-77%; the alleles are however on average shorter than those of pλg3 and λMS32 (Table 1) and should be more amenable to amplification by PCR. The flanking sequences of pλg3, 33.1, 33.4 and 33.6 have been described previously [5,8]; the flanking DNA of λMS32 and pMS51 was sequenced as described before [8]. All flanking DNA sequences were screened against the EMBL DNA sequence database to identify repeat elements such as Alu, and PCR oligonucleotide primers A and B (Fig.1) were designed to avoid such elements. Details of all primers and hybridization probes are given in Fig.1 legend.

Fidelity and efficiency of PCR amplification of human minisatellite alleles

To determine the ability of Taq polymerase to amplify long minisatel-lite alleles in particular, a mixture of 0.1µg genomic DNA from each of 4 individuals, giving a total of 8 different λMS32 alleles ranging in length

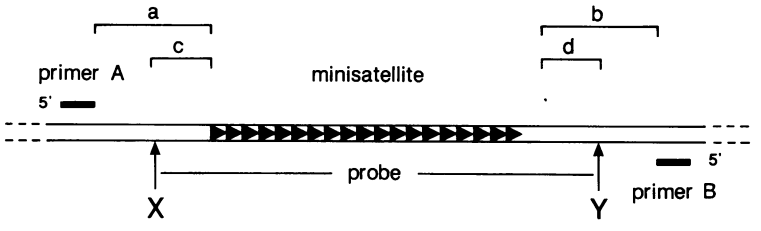

Figure 1. Primers and hybridization probes used in the amplification of minisatellites by PCR. Each minisatellite locus was amplified using 20- or 24-mer primers A and B located in unique sequence flanking DNA a and b bp respectively from the minisatellite. PCR products were detected by hybridization with an internal minisatellite probe isolated by cleavage with restriction endonucleases X and Y which cleave c and d bp from the minisatellite. Details for the six minisatellites are as follows, where R= length of repeat unit (bp): pλg3, A =5'-ACCACAGGCAGAGTAAGAGG-3', B=5'-CCACCCT GCTTACAGCAATG-3', X=PstI, Y=DdeI, a=35, b=58, c=26, d=45, R=37; λMS32, A=5'-TCACCGGTGAATTCCACAGACACT-3', B=5'-AAGCTCTCCATTTCCAGTTTCTGG-3', X=HpaI, Y=BglI, a=181, b=324, c=10, d=37, R=29; pMS51, A=5'-GATCAGCGAACTTCCTCTCGGCT C-3', B=5'-TCCACATTGAGGACTGTGGGAAGC-3', X=DdeI, Y=HaeIII, a=117, b=131, c=30, d=0, R=25 and 33; p33.1, A=5'-CTTTCTCCACGGATGGGATGCCAC-3', B=5'-GCCGTGTCAC CCACAAGCTTCTGG-3', X=DdeI, Y=RsaI, a=6, b=27, c=0, d=15, R=62; p33.4, A=5'-CC GGGCCAGACCCCAGCTGCTGAG-3', B= 5'-GCAGCATAGGGGCTGTCCTGG GCT-3', X=DdeI, Y=DdeI, a=11, b=96, c= 0, d=2, R=64; p33.6, A=5'-TG TGAGTAGAGGAGACCTCACATT-3', B=5'-AGGTGAGACATTACTCAATCCAAG-3',X=StyI, Y=DraIII, a=14, b=45, c= 10, d=16, R=37.

from 1.1 to 17.9 kb, was amplified for 10-20 cycles using λMS32 flanking primers A and B, followed by Southern blot hybridization with a minisatellite probe (Fig.2A). Using 6 min extension times for Taq polymerase, only the four shortest alleles (1.1-2.9 kb) were efficiently amplified. Increasing the extension time to 15 min, to improve the chance that the Taq polymerase would progress completely across the minisatellite, gave a marked increase in yield of the next two larger alleles (4.5, 6.6 kb), though no further improvement was seen with 30 min extensions. The relative yield of large alleles could also be improved by increasing the concentration of Taq polymerase (Fig.2B), allowing the detection of an allele, albeit faintly, 10.2 kb long. Addition of extra Taq polymerase at the 13th cycle gave only a marginal improvement in yield, and there is no evidence for a significant drop in polymerase activity during these prolonged extension times. Further experiments varying annealing temperature, extension temperature and buffer concentration failed to improve the yield of large alleles (data not shown), and all further experiments used 15 min extension times and high concentrations of Taq polymerase (1.5 units per 10μl PCR reaction).

Table 1. Properties of human minisatellites selected for PCR amplification.

Clone	Locus	Chromosome localisation	Heterozygosity (%)	No. alleles	Allelic length range (kb)[1]	%GC[2]	Ref.
pλg3	D7S22	7q36-qter	97	>40	0.6-20	66	[8,10,32]
λMS32	D1S8	1q42-q43	97	>40	1.1-20	62	[10,32]
pMS51	D11S97	11q13	77	9	1.3-4.3	69	[33]
p33.1	-	-	66	10	1.1-2.5	56	[5]
p33.4	-	-	70	7	0.8-1.3	68	[5]
p33.6	-	-	67	8	0.5-1.0	70	[5]

1; lengths of alleles include both the minisatellite and the flanking DNA defined by the PCR primers (see Fig. 1). 2; GC content of the minisatellite repeat units. The pMS51 minisatellite repeat unit has not been previously described, and is 5'-ACATGGCAGG(AGGGCAGG)$_n$TGGAGGG-3', where n=1 or 2 depending on the repeat unit.

At low cycle numbers (10 cycles), the alleles amplified appear to be completely faithful copies of the starting λMS32 alleles, as judged by their electrophoretic mobilities (Fig.2A). At higher cycle numbers (14, 17 cycles), there is an increase in background labelling; since most of this can be eliminated by digestion with S1 nuclease (data not shown), much of this background probably arises from low levels of single-stranded templates from the previous cycle which have failed to prime, and from incomplete extension products from the previous cycles which by definition cannot prime. At high cycle numbers (20), the hybridization pattern degenerates to a heterodisperse smear, as expected since the yield of PCR product becomes so high (>400ng/ml) that out-of-register annealing of single-stranded tandem-repeated minisatellite DNA will occur during the extension phase. This will lead to the premature termination of extension at a reannealed site, to spurious "alleles" arising from the extension of incomplete templates annealed out-of-register to the complementary strand of a minisatellite, and to the formation of multimolecular networks of reannealed minisatellite DNA strands. As a consequence, it is not yet possible to amplify minisatellites faithfully to the point where alleles can be visualised directly on ethidium bromide-stained agarose gels (data not shown).

The yields of each λMS32 allele amplified by PCR were quantified by scanning densitometry (Fig.3). PCR products from 0.1μg genomic DNA accumulate exponentially at least up to cycle 17. The gain in product per cycle decreases monotonously with allele length, with lower gains for 6 min compared with 15 min extension times. The gain versus allele length curves

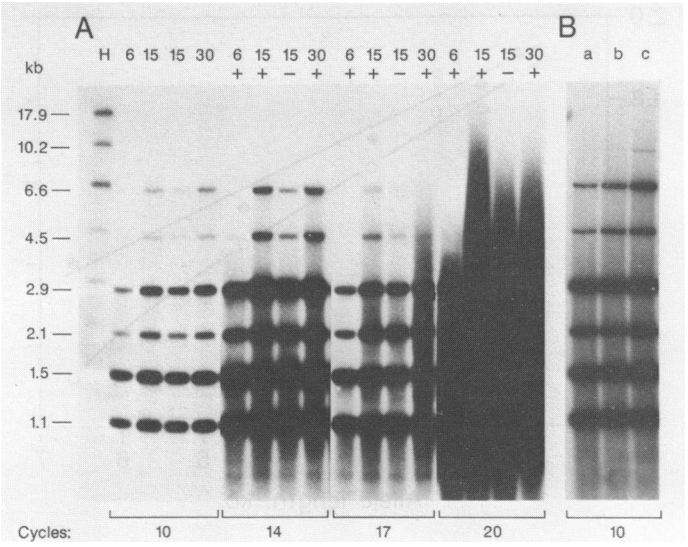

Figure 2. Amplification of λMS32 minisatellite alleles by PCR. A, 0.1μg aliquots of DNA from CEPH individuals 2306, 10208, 133101 and 133304, which together contain 8 different λMS32 alleles ranging from 1.1 to 17.9 kb, were pooled and amplified for 10-20 cycles in 10μl reactions containing 1 unit Taq polymerase plus flanking primers A and B. PCR products were separated by electrophoresis in a 1% agarose gel and detected by Southern blot hybridization with a minisatellite probe. Taq polymerase extension times at 70° were for 6, 15 or 30 min, with (+) or without (-) addition of extra polymerase (1 unit) at the 10th cycle. H, 2μg of each CEPH DNA digested with AluI; AluI sites flanking λMS32 are located such that each AluI allele is 0.2 kb longer than its corresponding PCR product. Autoradiography was for 5 hr (cycles 10, 14) or 1 hr (17, 20) without an intensifier screen. B, effect of increasing concentration of Taq polymerase (a-c, 0.5, 1, 2 units respectively) on the efficiency of amplification of large alleles. The extension time at 70° was 15 min.

extrapolate back to a gain per cycle of approximately 2.0 for very short alleles, indicating that the efficiency of denaturation and priming at each cycle is close to 100%. Final yields of an allele can be calculated from these curves; for an allele A with gain g_A per cycle present initially at n molecules, the yield after c cycles is approximately $n.g_A^c$ molecules. The molar imbalance between alleles A and B of different lengths, arising through more efficient amplification of shorter alleles, is given by $(g_A/g_B)^c$. For example, after 10 cycles of amplification with 15 min extension times, the molar yield of a 1 kb allele will be 18 times higher than that of a 6 kb allele; after 25 cycles, the imbalance will be 1300-fold. This imbalance is diminished to some extent by the more efficient detection of longer alleles

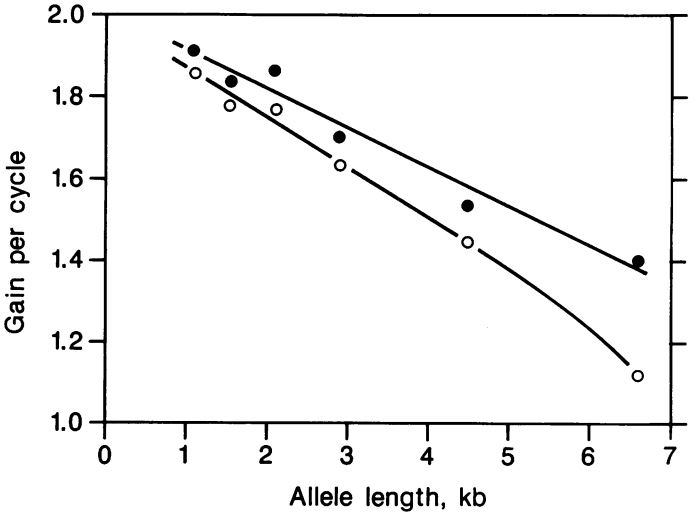

Figure 3. Efficiency of amplification of λMS32 minisatellite alleles as a function of allele length, with PCR extension times of 6 min (O) or 15 min (●). The gain in product per amplification cycle was determined by scanning laser densitometry of tracks H, 6+ and 15+ of Fig.3, exposed to pre-flashed X-ray film without an intensifier screen. The mean estimates of gain per cycle determined up to cycle 10, from cycle 10 to 14 and from cycle 14 to 17 were in close agreement, indicating that the yield of PCR product is increasing exponentially at least up to cycle 17; the Figure shows the mean value of the three estimates of the gain for each allele.

by the minisatellite hybridization probe. Nevertheless, long alleles amplified by PCR will become increasingly difficult to detect with the high numbers of PCR cycles needed to analyse very small amounts of starting human genomic DNA.

Minisatellites pλg3, pMS51, 33.1, 33.4 and 33.6 were also tested for their ability to be amplified by PCR (data not shown). In all cases, faithful amplification of all alleles tested was observed, except for the longest (>8 kb) alleles of pλg3 which as expected failed to amplify. Again, yields of PCR product fell with increasing allele length.

Fidelity of amplification of single minisatellite molecules

To test whether faithful amplification of single molecules is possible, 6 and 60 pg aliquots of human DNA, equivalent to 1 and 10 cells respectively, were co-amplified for 25 cycles using primers for both λMS32 and pMS51 (Fig. 4A). Both alleles of pMS51 (1.6 and 1.5 kb) amplified in the 60pg sample, and amplification products of one or both alleles were also seen in most of the 6pg samples, indicating that single target molecules can be faithfully

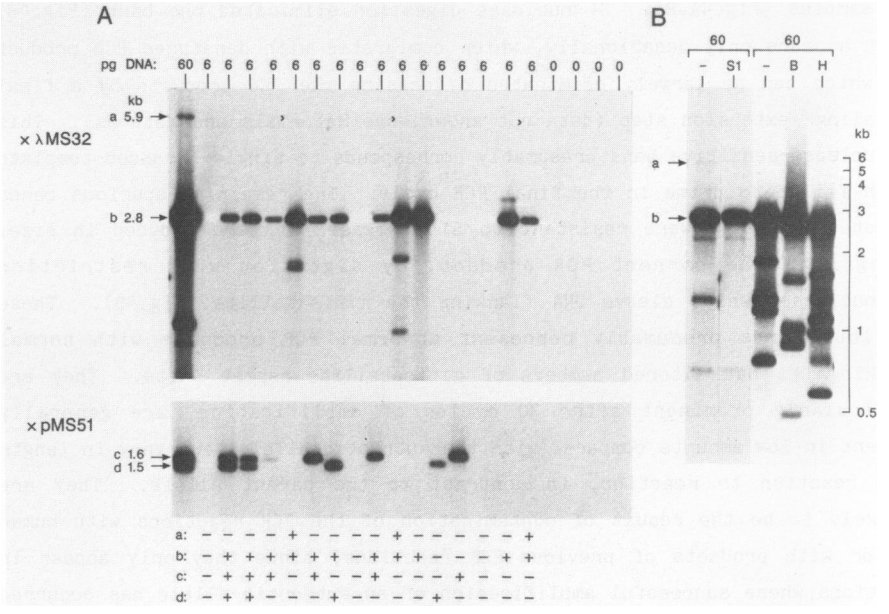

Figure 4. Co-amplification of two minisatellites from single cell equivalents of human DNA. A, 60 or 6pg aliquots of DNA from blood from an individual heterozygous for alleles a, b at λMS32 and c, d at pMS51 were amplified for 25 cycles with 15 min extension times in the presence of primers A and B for both loci, followed by Southern blot hybridization analysis of amplification products. Low levels of allele a could be detected in three of the 6pg samples on prolonged autoradiographic exposure (arrows). B, analysis of spurious amplification products of λMS32. Two 60pg aliquots of DNA were amplified for 30 cycles, followed by digestion with S1 nuclease (S1), BglI (B) or HpaI (H). BglI cleaves once in the flanking DNA, between the λMS32 minisatellite and primer B, and removes 311bp of flanking DNA. HpaI cleaves between primer A and the minisatellite, removing 195bp of flanking DNA (see Fig 1. legend).

amplified. Similarly, the 2.8 and 5.9 kb alleles of λMS32 could be successfully amplified from 6pg samples of human DNA, although the yield of the larger allele was as expected very low. Successful amplification of λMS32 and pMS51 alleles in 6pg samples appeared to occur independently, as expected, with a mean failure rate per allele per reaction of 63%. From the Poisson distribution, this indicates on average 0.46 successful amplification events per 6pg DNA sample,compared with 1 event predicted since 6pg human DNA will on average contain one molecule of an allele. Thus, single target minisatellite molecules can, with reasonable efficiency, be amplified by PCR.

No spurious amplification products were seen with pMS51. In contrast, λMS32 frequently gave unexpected products in both the 60pg and 6pg

DNA samples (Fig.4A,B). S1 nuclease digestion eliminated one band (Fig.4B) which appears only occasionally, which comigrates with denatured PCR product and which can be largely eliminated by chasing the PCR products by a final annealing/ extension step (data not shown, see Materials and Methods). This S1 nuclease-sensitive band presumably corresponds to single-stranded template which failed to prime in the final PCR cycle. The remaining spurious bands detected by λMS32 were resistant to S1 nuclease but were reduced in size, along with the correct PCR product, by digestion with restriction endonucleases which cleave DNA flanking the minisatellite (Fig.4B). These spurious bands presumably represent abnormal PCR products with normal flanking DNA but altered numbers of minisatellite repeat units. They are particularly prominent after 30 cycles of amplification, are generally present in low amounts compared with the authentic allele, and vary in length from reaction to reaction, in contrast to the parent allele. They are unlikely to be the result of contamination of the PCR reactions with human DNA or with products of previous PCR reactions, since they only appear in reactions where successful amplification of an authentic allele has occurred (Fig.4A) and have been consistently seen with all human DNAs tested (data not shown). Since almost all of the spurious products are shorter than the authentic allele, it is likely that they arise fairly early in the PCR reaction and accumulate preferentially due to their short length and concomitant higher efficiency of amplification. It is not yet clear how these "mutant" alleles arise, nor whether PCR conditions can be found which will suppress their apperance. A similar frequency of appearance of abnormal "alleles" has been seen with pλg3, and at a lower frequency with the other four minisatellites tested.

Somatic mutations at minisatellite loci in the starting human genomic DNA could also be a source of unexpected PCR products. Such mutations do exist, particularly for λMS32, as shown by the appearance of mutant minisatellite alleles in clonal tumour cell populations [30]. Somatic mutants are however unlikely to be a major source of the spurious bands shown in Fig.4, since no PCR reactions on 6pg human DNA have yet been seen which show a mutant allele appearing in the absence of the normal parental allele.

Co-amplification of multiple minisatellites: PCR-derived human DNA fingerprints

Fig.4 demonstrates that two minisatellites can be successfully co-amplified in the same PCR reaction. Further analyses showed that at least six minisatellites could be co-amplified without any apparent interference

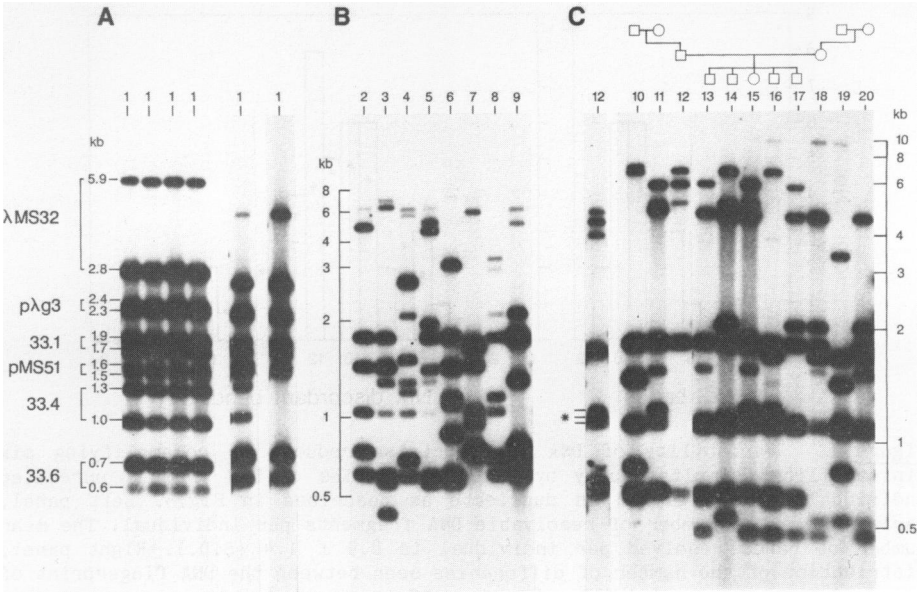

Figure 5. Co-amplification of six different human minisatellites by PCR. A, amplification of 10ng (first four lanes) or 1ng DNA (last two lanes) from individual 1 for 15 or 18 cycles respectively, using a cocktail of primers A and B for minisatellites p λg3, λMS32, pMS51, 33.1, 33.4 and 33.6. PCR products were detected by Southern blot hybridization with a cocktail of all six ^{32}P-labelled minisatellite probes. The individual tested had been previously characterised at all six loci separately, which enabled all hybridizing DNA fragments to be assigned as shown; this individual is heterozygous at all six loci. These DNA fingerprints are from three separate experiments. Note that 33.4 has failed to amplify in the last track. B, DNA fingerprints of 8 unrelated individuals (2-9) following amplification of 1ng samples of DNA for 18 cycles. C, DNA fingerprints of a 3-generation family (CEPH kindred 1435), following amplification of 10ng DNA for 15 cycles. Three bands, corresponding to alleles of 33.4 and pMS51, failed to amplify in individual 12, as shown by a second analysis of this family (first track, bands marked with an asterisk). In all experiments, PCR products were digested with S1 nuclease (see Materials and Methods) prior to gel electrophoresis, to reduce background labelling. DNA-free controls in all experiments were consistently blank (not shown).

between loci. Furthermore, the PCR products could also be typed simultaneously by Southern blot hybridization with a cocktail of all six minisatellite probes.

Examples of such multilocus PCR-derived DNA "fingerprints" are shown in Fig.5. In all cases, the PCR reaction was limited to 15-18 cycles, to minimise the appearance of spurious products as seen in Fig.4. These DNA

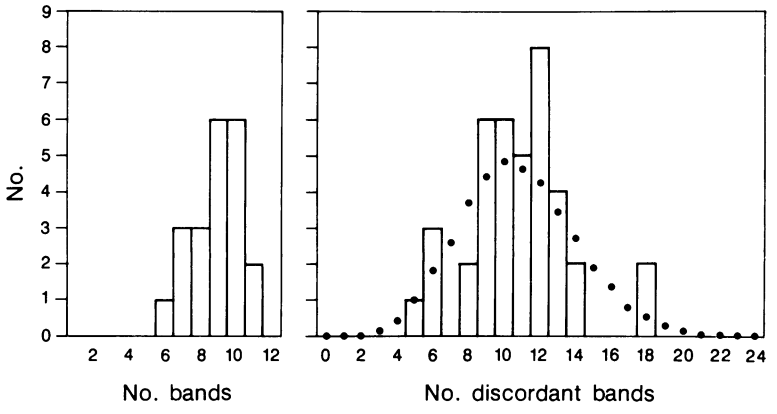

Figure 6. Variability of DNA fingerprints produced by co-amplifying six minisatellites simultaneously by PCR. 1ng samples of DNA from 21 unrelated individuals were analysed in duplicate as described in Fig.5. Left panel, variation in the number of resolvable DNA fragments per individual. The mean number of bands resolved per individual is 8.9 ± 1.4 (S.D.). Right panel, distribution of the number of differences seen between the DNA fingerprint of pairs of unrelated individuals, based on 39 independent pairwise comparisons. The number of discordant bands is the total number of bands not shared by the two individuals being compared. The theoretical maximum number of discordancies with 6 loci is 24, and the observed mean is 10.8 ± 2.8 (S.D.). The distribution of discordancies approximates to a Poisson distribution with this mean (dots).

"fingerprints" could be readily derived from 1ng human DNA. Repeat analyses of the same individual showed that the pattern was reproducible, with all hybridizing DNA fragments representing authentic minisatellite amplification products. On occasion, one or two loci failed to amplify (individuals 1,12, Fig.5A,C); this failure usually affected 33.4, followed by pMS51, and was least likely to affect 33.1 (data not shown). The likelihood of failure appears to correlate with the GC content of the minisatellite repeat units (Table 1), and suggest that non-amplification results from failure to denature GC-rich minisatellites at 95°, probably due to localised temperature variations in the heating block or to poor thermal conductivity between the block and the reaction tube.

These PCR DNA fingerprints are derived from six loci with widely differing levels of variability (Table 1). To determine the overall complexity and level of variability of these patterns, unrelated individuals were compared (Fig.5B). On average, 8.9 bands were resolved per individual (range 6-11, Fig.6). The maximum possible number of bands is 12 (Fig.5A), corresponding to heterozygosity at all loci, with no electrophoretic

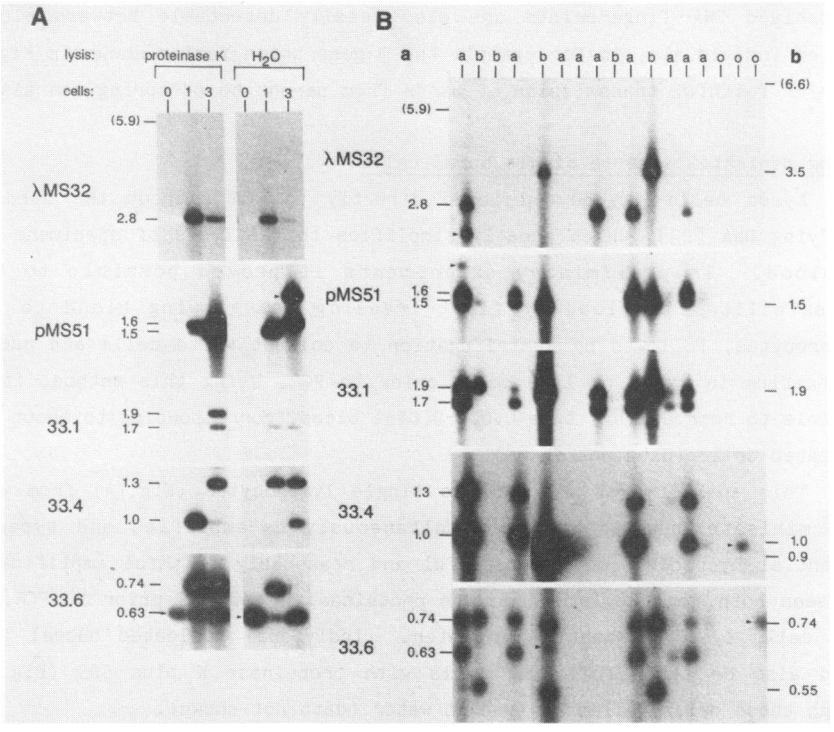

Figure 7. Amplification of minisatellites from single human cells. A, samples containing 0, 1 or 3 lymphocytes (from the individual analysed in Fig.5A) were lysed either with proteinase K plus SDS or by heating in water, followed by co-amplification with primers for λMS32, pMS51, 33.1, 33.4 and 33.6 for 27 cycles. PCR products were Southern blot hybridized sequentially with each of the five minisatellite probes. B, amplification products of single buccal cells, analysed following lysis with proteinase K plus SDS and PCR as above. Cells from two individuals, a and b were tested; b is homozygous at pMS51, 33.1 and 33.6. 0, no cell control. Spurious PCR products are indicated with arrows.

comigration of alleles from different loci and with no alleles too large to be amplified by PCR. In pairwise comparisons of unrelated individuals, there are on average 10.8 bands which are discordant between pairs of individuals (range 5-18). Since the distribution of discordancies approximates to a Poisson distribution, then the chance that two unrelated individuals would show identical DNA fingerprints (no discordant bands) can be estimated at $e^{-10.8} = 2 \times 10^{-5}$. These patterns therefore show a good degree of individual-specificity, despite the fact that four of the six loci used show relatively modest levels of variability (Table 1). Differences in

PCR-derived DNA fingerprints are also readily detectable between closely related individuals, in for example the 3-generation family shown in Fig.5C, in which faithful transmission of bands from parent to offspring can also be seen.

Typing minisatellites in single human cells

Lysed cells can be subjected directly to PCR without the need for purifying DNA [23], which greatly simplifies the analysis of specimens such as blood. In preliminary experiments it proved possible to type minisatellites in blood by first freezing and thawing blood to lyse erythrocytes, followed by centrifugation to collect white cells and nuclei, and heating in water to lyse cells prior to PCR. Using this method, it was possible to reproducibly type 0.001-0.01μl blood, corresponding to about 5-50 nucleated cells (data not shown).

This analysis was extended to single lymphocytes (Fig.7A) from which five minisatellite loci were simultaneously co-amplified and typed by sequential hybridization. Successful and reasonably faithful amplification was seen both from cells lysed with proteinase K and SDS prior to PCR, and from cells lysed by heating in water. Individual nucleated buccal cells could also be typed following lysis with proteinase K plus SDS (Fig.7B), though these cells failed to lyse in water (data not shown).

To test the feasibility of identifying individual cells, 14 buccal cells from two individuals were separately typed in a blinded experiment (Fig. 7B). In four cases, no amplification products were seen from any of the loci, suggesting either that the cell had not been transferred to the PCR reaction, or that lysis had not occurred, or that nuclear DNA had degraded prior to PCR. In the remaining 10 cases, amplified alleles could be detected from at least two of the minisatellite loci, and in some cases all five loci amplified successfully from a single cell. Omitting the large alleles of λMS32, which amplify poorly and would be difficult to type at the single cell level, we estimate that, for those single cell PCR reactions in which at least some loci have amplified, approximately 75% of alleles present could be detected following PCR. This estimate agrees with the efficiency of single molecule amplification determined from PCR analysis of 6pg samples of human DNA (Fig.4). As expected from Fig.4, several instances of spurious bands were seen in both buccal cell and lymphocyte PCR reactions (Fig.7A,B). Nevertheless, distinguishing alleles from each of the two individuals tested could be detected in the 10 successfully-typed buccal cells, and the origin of each buccal cell was successfully predicted in this blinded trial.

Finally, we note the presence of amplified products of 33.6 in some of the DNA-free controls in Fig.7A,B and of 33.4 in one of the buccal cell controls. In practice, we have found that such contamination, probably with recombinant DNAs or the products of previous PCR reactions rather than with human cells or genomic DNA, can only be avoided by using solutions, glassware, disposable pipette tips and microcentrifuge tubes which have not previously been exposed to the laboratory environment. It is noteworthy that our most consistent contamination problem has been with 33.6, one of the multilocus DNA fingerprint probes [5,6] which has been in continuous use in our laboratory for the last four years.

DISCUSSION

Taq polymerase not only shows remarkable fidelity in amplifying non-repeated DNA [21], but is also capable of faithfully amplifying entire minisatellites and preserving the allelic specificity of the number of repeat units. Unlike conventional PCR reactions, however, minisatellite PCR must be terminated before the yield of product becomes so high (>4ng per 10µl PCR reaction) that out-of-register annealing between complementary tandem-repeated template strands occurs during primer extension, particularly during the lengthy extension times needed to obtain efficient amplification of long minisatellite alleles. Also, the PCR reaction must proceed far enough to generate sufficient product to be detectable by hybridization. The minisatellite probes are sensitive and can readily detect 0.1pg minisatellite PCR product [10]. The "window" of PCR cycles which generate an appropriate amount of product for typing (0.1-4000 pg product) is therefore very wide, and only a very approximate estimate of the amount of initial human genomic DNA is needed to predict the number of PCR cycles required for successful typing. As a guide, 10-15 cycles are appropriate for 100ng genomic DNA in a 10µl PCR reaction, 18 cycles for 1ng DNA and 25 cycles for single cell PCR (6pg). The number of PCR cycles may need to be increased to detect larger alleles which amplify less efficiently.

Since minisatellite amplifications need to be restricted to the exponential phase of accumulation of PCR products [21], then the hybridization signal is approximately proportional to the amount of input DNA down to 0.1ng human genomic DNA (data not shown). Below this level, stochastic variation in the number of target minisatellite molecules can obscure the proportionality. Minisatellite PCR can therefore be used quantitatively to estimate low concentrations of human DNA. Also, the amount

of primers and Taq polymerase will not be limiting during this early phase of PCR, and in principle there should be little or no interference between different loci being amplified. In practice, at least six different minisatellites can be co-amplified simultaneously, and there seems to be no theoretical reason why this number could not be increased further.

Co-amplification of minisatellites followed by simultaneous or sequential hybridization with minisatellite probes enables a considerable amount of information concerning individual identity and family relationship to be gathered from very small DNA samples. These PCR-derived DNA fingerprints appear to be reliable down to 1ng human DNA. Information can also be recovered from much lower amounts of DNA and single cells, although the generation of spurious DNA fragments at some loci, during the relatively large number of PCR cycles needed for single cell typing, could present significant problems for individual identification at the level of one or a few cells. Fortunately, these spurious PCR products appear to vary from reaction to reaction, and duplicate PCR analyses of very small samples of DNA should therefore distinguish bona fide amplified alleles from spurious PCR products.

PCR-derived DNA fingerprints already show a good level of individual specificity, with a chance of false association of two individuals of approximately 2×10^{-5}. In contrast, conventional DNA fingerprints obtained by Southern blot hybridization with a multilocus polycore probe [6] or with a cocktail of locus-specific human minisatellite probes [10] show much higher levels of individual specificity ($<10^{-12}$ and $<10^{-6}$ respectively). However, several approaches could be used to improve the variability of PCR-derived DNA fingerprints. First, highly informative alleles particularly at pλg3 and λMS32 cannot be detected above approximately 8 kb (1ng human DNA) or approximately 5 kb (single cell). This could be overcome by using highly variable minisatellites with a more restricted range of allele lengths. Such loci appear to be scarce since high levels of variability are usually associated with large numbers of minisatellite repeat units and long alleles [5,8,10]. Some possibly appropriate loci have however been isolated ([9], J.A.L. Armour and A.J. Jeffreys, unpublished data). Second, the number of minisatellites being amplified simultaneously could be increased. Third, loci which are particularly prone to generate spurious PCR products, such as pλg3 and λMS32, could be identified and avoided. If these goals can be accomplished, then we see no reason why reliable identification at the single cell level should not be possible, provided that inadvertant contamination of

PCR reactions is avoided and that the potential presence of somatic mutations at hypervariable loci is taken into account [30].

The use of multilocus DNA fingerprint probes in for example individual identification in forensic medicine, paternity testing and monitoring bone marrow transplants is limited by the sensitivity of these probes which require at least 0.1-1μg human DNA for typing [6]. Similarly, locus-specific minisatellite probes can only be used successfully on a minimum of approximately 50ng human DNA [10]. PCR-derived DNA fingerprinting improves sensitivity by orders of magnitude and can be used to type specimens which are relatively intractable by conventional Southern blot hybridization. For example, human hair roots typically contain 10-500ng DNA [22] and while approximately 70% of roots can be typed using locus-specific minisatellite probes (Z. Wong, J.A.L. Armour and A.J. Jeffreys, unpublished data), all hair roots so far tested can be typed by PCR-derived DNA fingerprinting (data not shown). Similarly, 0.001-0.01μl blood can be typed without the need first to purify DNA. Likewise, saliva contains on average 400 nucleated buccal cells per μl (range 100-800 in 14 individuals tested), and PCR-derived DNA fingerprint analysis of submicrolitre samples of saliva is therefore possible. The potential for typing trace amounts of hair, blood, semen, saliva and urine in forensic specimens, including partially degraded samples, is obvious. The potential for inadvertant contamination of specimens, for example with traces of saliva, is likewise evident.

PCR-derived DNA fingerprints should eventually become sufficiently individual-specific to provide a statistically highly polarised test for establishing parentage in for example paternity disputes. Not only would the need to isolate DNA be obviated, but much smaller samples of blood obtained by finger-pricking rather than venepuncture could be used. Alternatively, the determination of parentage could be based on the analysis of saliva. This would avoid the problem of individuals who object to giving blood samples, and would remove the trauma of taking blood from infants.

Finally, PCR could be used to study mutation processes at minisatellites. The minisatellites used in the study, and the multilocus DNA fingerprint probes, share a common "core" sequence in the repeat units which we suspect may serve as a recombination signal promoting unequal crossing over at these tandem-repeated loci and generating high levels of allelic variability [4,5]. A significant level of germline and somatic instability at the most variable human minisatellites has been detected by studying human pedigrees [31] and clonal human tumours [30] respectively. The ability to

amplify single target minisatellite molecules will permit a much more detailed analysis of the rate of production of mutant minisatellite alleles and the structure of the products of the (recombinational?) mutation processes which generate these hypervariable loci.

ACKNOWLEDGEMENTS

We are grateful to Professors J.Dausset and H.Cann for providing CEPH DNAs, and to Martin Woodhouse and other volunteers who provided specimens. We thank Raymond Dalgleish, Karen Molyneux, John Armour, Andrew Collick, Robert Kelly and Nicola Royle for helpful discussions. This work was supported by grants from the Medical Research Council. A.J.J. is a Lister Institute Research Fellow. The minisatellite probes and the work described herein are the subjects of Patent Applications, and commercial enquiries should be addressed to ICI Diagnostics, Gadbrook Park, Northwich, Cheshire, CW9 7RA, U.K.

*To whom correspondence should be addressed

REFERENCES

1. Bell,G.I., Selby,M.J. and Rutter,W.J. (1982). Nature 295, 31-35.
2. Capon,D.J., Chen,E.Y., Levinson,A.D., Seeburg,P.H. and Goeddel,D.V. (1983). Nature 302, 33-37.
3. Goodbourn,S.E.Y., Higgs,D.R., Clegg,J.B. and Weatherall,D.J. (1983). Proc.Nat.Acad.Sci. USA. 80, 5022-5026.
4. Jeffreys,A.J. (1987). Biochem.Soc.Trans. 15, 309-317.
5. Jeffreys,A.J., Wilson,V. and Thein,S.L. (1985). Nature 314, 67-73.
6. Jeffreys,A.J., Wilson,V. and Thein,S.L. (1985). Nature 316, 76-79.
7. Vassart,G., Georges,M., Monsieur,R., Brocas,H., Lequarre,A.S. and Christophe,D. (1987). Science 235, 683-684.
8. Wong,Z., Wilson,V., Jeffreys,A.J. and Thien,S.L. (1986). Nucleic Acids Res. 14, 4605-4616.
9. Nakamura,Y., Leppert,M., O'Connell,P., Wolff,R., Holm,T., Culver,M., Martin,C., Fujimoto,E., Hoff,M., Kumlin,E. and White,R. (1987) Science 235, 1616-1622.
10. Wong,Z., Wilson,V., Patel,I., Povey,S. and Jeffreys,A.J. (1987). Ann.Hum.Genet. 51, 269-288.
11. Reeders,S.T., Breuning,M.H., Davies,K.E., Nicholls,R.D., Jarman,A.P., Higgs,D.R., Pearson,P.L. and Weatherall,D.J. (1985). Nature 317, 542-544.
12. Donis-Keller,H., Green,P., Helms,C., Cartinhour,S., Weiffenbach,B., Stephens,K., Keith,T.P., Bowden,D.W., Smith,D.R., Lander,E.S., Botstein,D., Akots,G., Rediker,K.S., Gravius,T., Brown,V.A., Rising, M.B.,Parker,C., Bricker,A., Phipps,P., Muller-Kahle,H., Fulton,T.R., Ng,S., Schumm,J.W., Barker, D.F., Crooks,S.M., Lincoln,S.E., Daly,M.J. and Abrahamson,J. (1987). Cell 51, 319-337.
13. Nakamura,Y., Lathrop,M., O'Connell,P., Leppert,M., Barker,D., Wright,E., Skolnick,M., Kondoleon,S., Litt,M., Lalouel,J.-M. and White,R. (1988). Genomics 2, 302-309.
14. Jeffreys,A.J., Brookfield,J.F.Y. and Semeonoff,R. (1985). Nature 317,818-819.

15. Baird,M., Balazs,I., Giusti,A., Miyazaki,L., Nicholas,L., Wexler,K., Kanter,E., Glassberg,J., Allen,F., Rubinstein,P. and Sussman,L. (1986). Am.J.Hum.Genet. 39, 489-501.
16. Thein,S.L., Jeffreys,A.J. and Blacklock,H.A. (1986). Lancet ii, 37.
17. Knowlton,R.G., Brown,V.A., Braman,J.C., Barker,D., Schumm,J.W., Murray,C., Takvorian,T., Ritz,J. and Donnis-Keller,H. (1986). Blood 68, 378-385.
18. Gill,P., Jeffreys,A.J. and Werrett,D.J. (1985). Nature 318, 577-579.
19. Gill,P., Lygo,J.E., Fowler,S.J. and Werrett,D.J. (1987). Electrophoresis 8, 38-44.
20. Mullis,K., Faloona,F., Scharf,S., Saiki,R., Horn,G. and Erlich,H. (1986). Cold Spring Harbor Symp.Quant.Biol. 51, 263-273.
21. Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J., Higuchi,R., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1988). Science 239, 487-491.
22. Higuchi,R., Von Beroldingen,C.H., Sensabaugh,G.F. and Erlich,H.A. (1988). Nature 332, 543-546.
23. Li,H., Gyllensten,U.B., Cui,X., Saiki,R.K., Erlich,H.A. and Arnheim,N. (1988). Nature 355, 414-417.
24. Wong,C., Dowling,C.E., Saiki,R.K., Higuchi,R.G., Erlich,H.A. and Kazazian,H.H. (1987). Nature 330, 384-386.
25. Saiki,R.D., Scharf,S.J., Faloona,F., Mullis,K.B., Horn,G.T., Erlich,H.A. and Arnheim,N. (1985). Science 230, 1350-1354.
26. Jeffreys,A.J. and Morton,D.B. (1987). Animal Genetics 18, 1-15.
27. Vieira,J. and Messing,J. (1982). Gene 19, 259-268.
28. Feinberg,A.P. and Vogelstein,B. (1984). Anal.Biochem. 137, 266-267.
29. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982). Molecular cloning: a laboratory manual (Cold Spring Harbor Laboratory, N.Y.).
30. Armour,J.A.L., Patel,I., Thein,S.L., Fey,M.F and Jeffreys,A.J. (1988). Manuscript submitted.
31. Jeffreys,A.J., Royle,N.J., Wilson,V. and Wong,Z. (1988). Nature 332, 278-281.
32. Royle,N.J., Clarkson,R.E., Wong,Z. and Jeffreys,A.J. (1988). Genomics, in press.
33. Royle,N.J., Clarkson,R., Wong,Z. and Jeffreys,A.J. (1987) Human gene mapping 9: ninth international workshop on human gene mapping. Cytogenet. Cell Genet. 46, 685.