Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 1946–1949, March 1994
Biochemistry

Sequence-specific "gene signatures" can be obtained by PCR with single specific primers at low stringency

SÉRGIO D. J. PENA*†‡, GUILLERMO BARRETO†, ANNAMARIA R. VAGO†, LUIZ DE MARCO§, FERNANDO C. REINACH¶, EMMANUEL DIAS NETO‖, AND ANDREW J. G. SIMPSON‖

*Núcleo de Genética Médica de Minas Gerais (GENE/MG), Avenue Afonso Pena 3111/9, Belo Horizonte, Brazil 30130-909; Departments of †Biochemistry and §Pharmacology, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil 31270-901; ¶Department of Biochemistry, Instituto de Química, Universidade de São Paulo, São Paulo, Brazil 01498; and ‖Centro de Pesquisas "René Rachou," Fundação Oswaldo Cruz, Belo Horizonte, Brazil 31090-002

Communicated by Charles R. Cantor, November 10, 1993

ABSTRACT **Low-stringency single specific primer PCR (LSSP-PCR) is an extremely simple PCR-based technique that detects single or multiple mutations in gene-sized DNA fragments. A purified DNA fragment is subjected to PCR using high concentrations of a single specific oligonucleotide primer, large amounts of *Taq* polymerase, and a very low annealing temperature. Under these conditions the primer hybridizes specifically to its complementary region and nonspecifically to multiple sites within the fragment, in a sequence-dependent manner, producing a heterogeneous set of reaction products resolvable by electrophoresis. The complex banding pattern obtained is significantly altered by even a single-base change and thus constitutes a unique "gene signature." Therefore LSSP-PCR will have almost unlimited application in all fields of genetics and molecular medicine where rapid and sensitive detection of mutations and sequence variations is important. The usefulness of LSSP-PCR is illustrated by applications in the study of mutants of smooth muscle myosin light chain, analysis of a family with X-linked nephrogenic diabetes insipidus, and identity testing using human mitochondrial DNA.**

In the past few years there has been extraordinary interest in the development of techniques capable of detecting DNA sequence variation. Major applications include the study of genetic disease, cancer, and identity testing. While DNA sequencing is obviously the most sensitive and informative method, it is too cumbersome for routine use in searching for mutations, especially when the DNA fragment of interest is large. For the detection of single-base mutations, the alternatives to sequencing consist of techniques dependent on prior knowledge of the exact base alterations [for instance, allele-specific hybridization (1, 2)] or shotgun strategies that ideally should detect any change [e.g., single-strand conformational polymorphism (3) and denaturing gradient gel electrophoresis (4)]. The former become awkward when one wishes to study sequence differences caused by a variable number of mutations (5), whereas the latter can only be applied to small DNA fragments and do not identify the number of modifications or their sites. We here report a technique called low-stringency single specific primer PCR (LSSP-PCR) that when applied to gene-sized DNA fragments (at least up to 1 kb) translates the underlying DNA sequence into a unique multiband "gene signature." Changes as small as single-base mutations significantly alter the signature, producing a phenotype that is diagnostic of the specific alteration. The usefulness of LSSP-PCR is illustrated by applications in the study of mutants of smooth muscle myosin light chain, analysis of a family with X-linked nephrogenic diabetes insipidus, and identity testing using the control (D-loop) region of human mtDNA.

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.

MATERIALS AND METHODS

Template DNA for LSSP-PCR. Template DNA was either in the form of purified plasmid inserts or DNA amplified by specific PCR. For the preparation of purified inserts, plasmids containing a 569-bp fragment of either normal or mutagenized chicken smooth muscle myosin light chain cDNAs were digested with *Bam*HI and *Eco*RI, the products were purified by low-melting-point agarose gel electrophoresis, and the concentration was estimated by comparison against known standards. For the amplification of the vasopressin V2 receptor, 100 ng of genomic DNA, isolated from peripheral blood by digestion with SDS/proteinase K followed by phenol/chloroform extraction and isopropanol precipitation (6), was subjected to 35 cycles of specific PCR. The annealing temperature used was 64°C for 1 min, and the denaturation and extension temperatures were, respectively, 94°C and 72°C. The primers used were NDI-1 (5′-TCTGCAGATGGTGGGCATGT-3′) and NDI-2 (5′-CTAGGTTGGGTTCAGGATG-3′), and the PCR conditions were as follows: each 20 μl of reaction mixture contained 2 units of *Taq* polymerase (Cenbiot, Rio Grande do Sul, Brazil), 200 μM dNTPs, 10 mM Tris·HCl (pH 8.6), 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, and 20 pmol of each primer. Fifteen microliters of reaction product was run on a 1% low-melting-point agarose gel, the gel was stained with ethidium bromide, and the specific amplified band was cut out. For the preparation of a 1021-bp region of the mtDNA control region, 200 ng of genomic DNA was amplified using the primers L15996 and H408 (7) and 0.25 unit of *Taq* polymerase. The reaction mixture consisted of 5 pmol of each primer, 50 μM dNTPs, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-100, and 10 mM Tris·HCl (pH 9.0) in a final volume of 25 μl. The cycle consisted of denaturation at 94°C for 45 sec, annealing at 56°C for 1 min, and extension at 74°C for 1 min. In the first cycle the denaturation was prolonged for 5 min. Ten microliters of reaction product was run on an 0.8% low-melting-point agarose gel, the gel was stained with ethidium bromide, and the amplified band was purified by excision from the gel.

LSSP-PCR. For the production of gene signatures from the chicken myosin cDNAs by LSSP-PCR,** 120 pmol of primer 5′-TTCGCTATGTTCGACCAGTCA-3′ (which represents nucleotides 82–102 of the cDNAs) was used together with 200 μM dNTPs, 4.0 units of *Taq* polymerase (Cenbiot), 10 mM Tris·HCl (pH 8.6), 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, and 15 pg of DNA in 25-μl reaction mixtures. Amplification was achieved using 35 repetitions of a two-phase cycle consisting of 1 min at 94°C and 1 min at 30°C. In the initial cycle, the denaturation was prolonged for 5 min. Fifteen microliters of each reaction was analyzed by poly-

Abbreviation: LSSP-PCR, low-stringency single specific primer PCR.

‡To whom reprint requests should be sent at the * address.

**This technique is the subject of a patent application.

acrylamide gel electrophoresis and silver staining (8). When PCR products were used for the production of gene signatures, the same protocol was used with one of the two specific primers used for the initial amplification as shown below. In the case of the vasopressin gene, we used the primer NDI-2 (5′-CTAGGTTGGGTTCAGGATG-3′) and 15 ng of template, whereas for mtDNA we used primer L15996 and 10-μl reaction volumes together with 15 ng of template and 48 pmol of primer. The production of fluorescent mtDNA gene signatures was achieved using fluorescein-labeled L15996 primer in a 25-μl reaction volume of which 10 μl was analyzed by gel electrophoresis and 5 μl was studied in an automatic laser fluorescence (A.L.F.; Pharmacia) using a 6-hr run on a denaturing 6% urea/polyacrylamide gel and default run conditions.

RESULTS

Mutants of Chicken Smooth Muscle Myosin Light Chain. The sensitivity and reproducibility of the procedure were explored by using three variants of a 569-bp cDNA fragment of chicken smooth muscle myosin light chain produced by *in vitro* mutagenesis (9) that differed by a single base from wild-type DNA as shown in Fig. 1*A*. The fragments were subjected to 35 cycles of LSSP-PCR involving a low annealing temperature (30°C), a high concentration of a single 20-bp specific primer (4.8 μM instead of the usual range of 0.1–1 μM), and elevated concentrations of *Taq* polymerase (160 units/ml) together with a standard reaction mixture. The amplification products were analyzed in polyacrylamide gels stained with silver salts. A complex pattern consisting of many bands was obtained, which was different with each mutant (Fig. 1*B*). The signatures were not unduly sensitive to the concentration of template, since the reproducibility of signatures obtained using up to a 5-fold variation was good (Fig. 1*C*). Moreover, when the procedure was undertaken on different days and with different template concentrations, but run in the same gel, we were able to detect only quantitative alterations in intensity of individual bands and not qualitative loss or gain of components of the signature (Fig. 1*D*). Thus, LSSP-PCR is reproducible. When we attempted to decrease the concentrations of *Taq* polymerase and primers, useful signatures were no longer obtained.

Vasopressin-Resistant Diabetes Insipidus. We have found that PCR products are also suitable templates for LSSP-PCR. Thus, starting with genomic DNA, we undertook a first step-specific PCR amplification of a region of interest, followed by LSSP-PCR using one of the two primers utilized for the initial amplification. We have applied the approach to a number of model systems. In Fig. 2 the results of a test case involving a family with X-linked nephrogenic diabetes insipidus, a disease caused by genetic defects in the vasopressin V2 receptor gene (10–12), are shown. We first amplified a 649-bp fragment of exon 2 of the vasopressin V2 receptor gene from total genomic DNA. Previous sequencing had shown that this fragment contained a single-base change in nucleotide 839 of the cDNA sequence (13) (A 839 → G, leading to a Tyr-280 → Cys substitution), which was responsible for the disease in this family (14). The 649-bp products of the initial specific PCR amplification of the proband, who presented with nephrogenic diabetes insipidus and had a family history of the disease segregating in a characteristic X-linked fashion, his carrier mother, and his normal father were gel-purified and subjected to LSSP-PCR (Fig. 2). The proband and the father both have a single copy of the gene, which, although differing by only one base substitution, produced clearly different signatures. In this instance, the major distinction resided in one prominent band of ≈310 bp (see the arrow in Fig. 2). The mother is heterozygous for the mutation and correspondingly has both of the bands that distinguished the father and son.

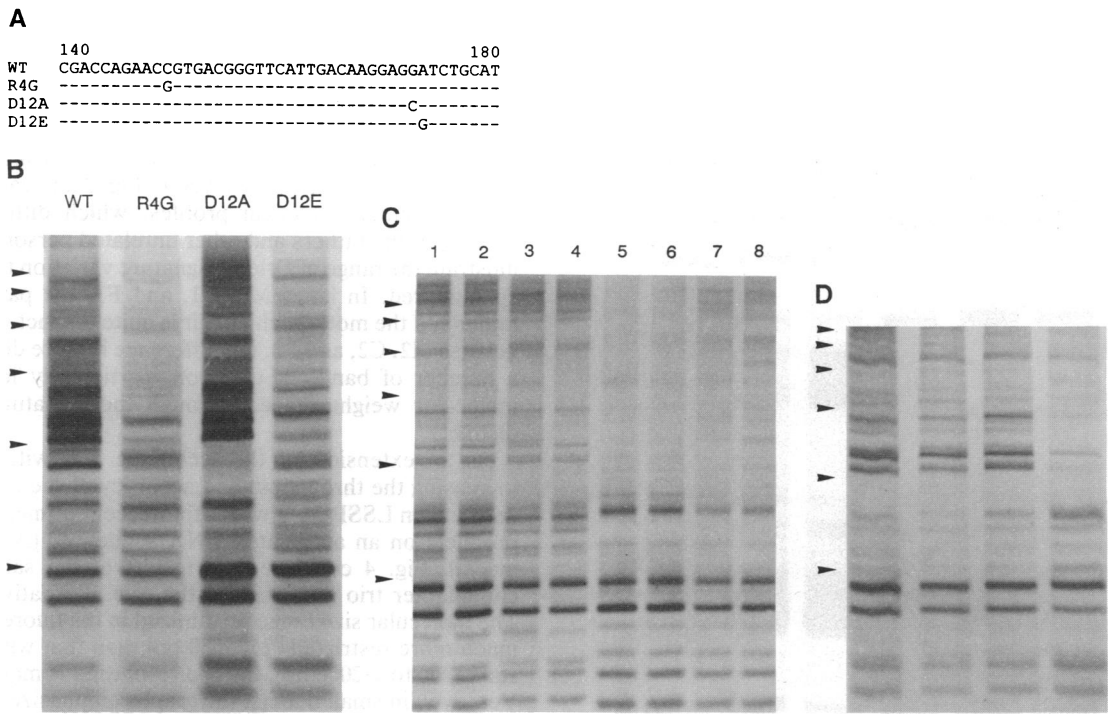

FIG. 1. (*A*) Sequence changes in 569-bp inserts of four cDNA clones containing a normal or mutagenized fragment of chicken smooth muscle myosin light chain. (*B*) An 8% polyacrylamide gel stained with silver salts exhibiting the LSSP-PCR signatures obtained with the four cDNAs. WT, wild type. (*C*) Silver-stained 8% polyacrylamide gel with duplicate signatures derived from R4G and D12A produced with different concentrations of template. Mutant D12A: lanes 1 and 2, 70 pg; lanes 3 and 4, 14 pg. Mutant R4G: lanes 5 and 6, 36 pg; lanes 7 and 8, 7.3 pg. (*D*) LSSP-PCR phenotypes of mutant D12A produced with different template concentrations and on different days. Lanes 1 and 4, day 1; lanes 2 and 3, day 2; lanes 1 and 2, 10 pg; lane 3, 20 pg; lane 4, 100 pg of template. The arrowheads on the left represent the positions of the 123-bp ladder size standards (123, 246, 369, 492, 611, and 738 bp from bottom to top).

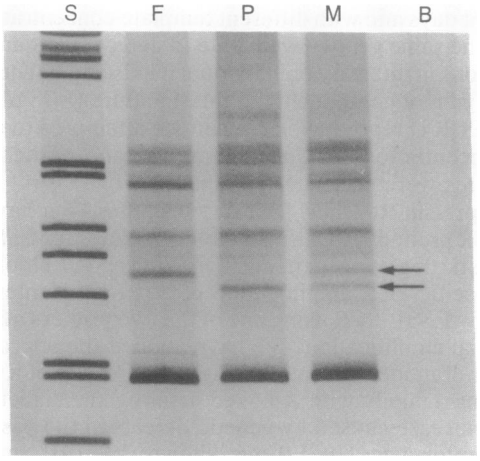

FIG. 2. Silver-stained 8% polyacrylamide gel showing signatures produced by amplifying the gene for the vasopressin V2 receptor from a proband (P), his carrier mother (M), and normal father (F). Lane S, molecular size standards (154, 203, 220, 298, 344, 396, 506, 517, 1018, 1636, and 2056 bp from bottom to top). Lane B, blank control without DNA.

Control (D-Loop) Region of Human mtDNA. Fig. 3 illustrates the data obtained with the more complex situation represented by the control (D-loop) region of human mtDNA, which is known to differ from between 1 and 15 bases between unrelated individuals (5, 14–16). In this instance we prepared human DNA samples from blood and used primers L15996 and H408 under specific conditions to amplify a 1021-bp portion of the mtDNA including the D-loop region.

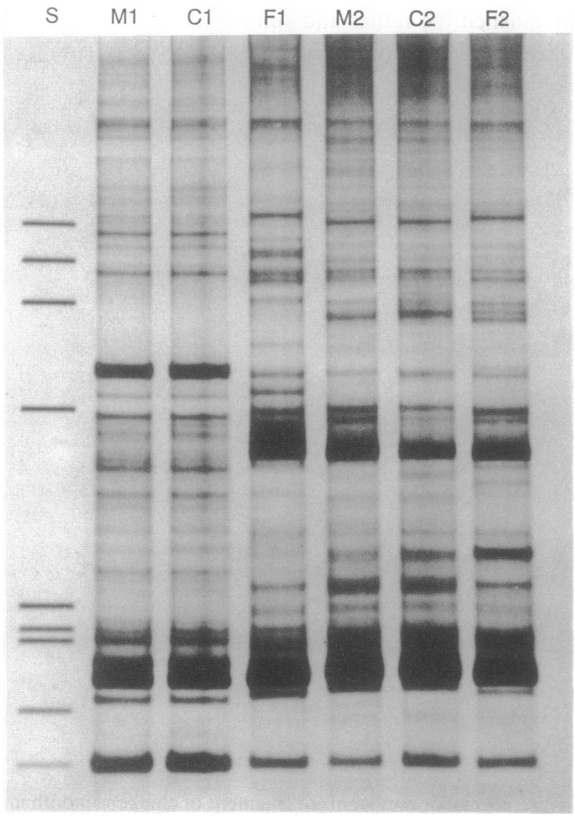

FIG. 3. Silver-stained 6% polyacrylamide gel showing signatures obtained using the 1021-bp D-loop region of mtDNA amplified from two trios composed of mother (M), child (C), and father (F). Lane S, size standards (118, 194, 234, 271, 281, 310, 603, 872, 1078, and 1353 bp from bottom to top).

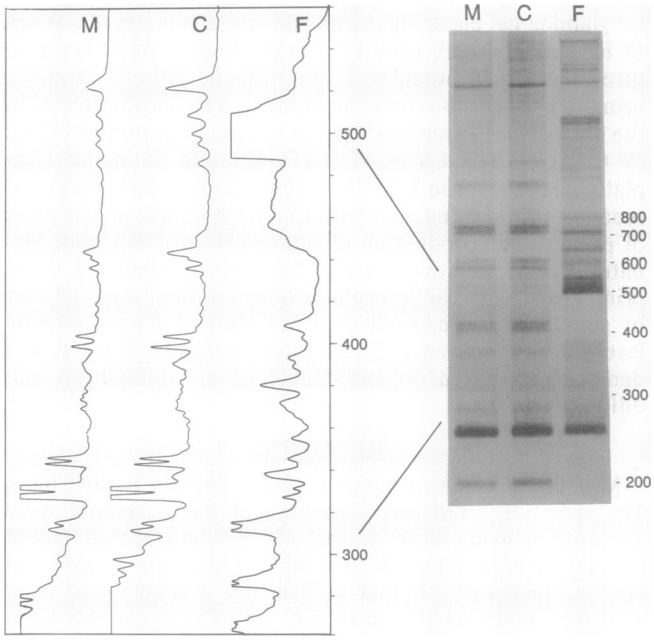

FIG. 4. Comparison of D-loop LSSP-PCR products resolved in silver-stained polyacrylamide gels (*Right*) and in an automatic laser fluorescence (A.L.F.; Pharmacia) sequencer (*Left*). Ten microliters of the reaction product was run on a 6% polyacrylamide gel and stained with silver (*Right*), and 5 μl was analyzed in the automated sequencer using a 6-hr run on a denaturing urea/6% polyacrylamide gel using the default conditions in the A.L.F. (*Left*). The three samples are a trio of mother (M), child (C), and father (F). The values on the right indicate size values in base pairs calculated from molecular weight standards.

The PCR product was gel purified and reamplified with the LSSP-PCR protocol using primer L15996 to produce complex D-loop signatures. Although dozens of bands are seen, the major definition of the signature depended on relatively few strong bands. We have compared 24 unrelated individuals and found different patterns in all of them (data not shown). In contrast, 12 mother–child pairs were identical, as expected from the matrilineal inheritance of mtDNA. Two mother–child pairs are displayed in Fig. 3 and can be seen to have absolutely identical profiles, which differ from the profiles of the fathers and other unrelated persons. The data illustrate the range of D-loop signature variation that we have encountered. In lanes M1, C1, and F1, the pattern of the father and the mother–child pair is quite distinct. In contrast, in lanes M2, C2, and F2, the difference is more discrete, with a number of bands in common, particularly in the lower molecular weight range, although the signatures are still distinct.

As an extension of the technique and with a view to increasing the throughput and accuracy of the test, we have undertaken LSSP-PCR with a fluorescent primer and run the products on an automated DNA sequencer (A.L.F.; Pharmacia). Fig. 4 compares the results of the same mother–child–father trio analyzed by the two alternative methods. The molecular size range scrutinized in the fluorescent run is much more restricted (250–550 bp) than that with the silver stain (200 to >2000 bp), but the resolution is much superior. Yet, even in small details, the graphs emphasize the identity of the mother's and child's mtDNA and their variation from that of the father.

DISCUSSION

In most instances when PCR is performed for preparative purposes, all precautions are taken to avoid hybridization of primers to regions other than those to which they are exactly

homologous. It has recently been appreciated, however, that hybridizations to mismatched sequences under conditions of low stringency can be useful. In particular, such an approach has led to the development of techniques based on randomly amplified polymorphic DNAs for the study of genetic polymorphisms. In such experiments, genomic DNA from an organism is used to generate the pattern of amplified DNA products (17, 18). In the present report, we describe a technique, which we have called LSSP-PCR, that differs from previous PCR approaches in two fundamental ways: a single *specific* primer is used under low-stringency conditions and, more importantly, PCR is applied to a *previously purified* DNA fragment, thus being transformed into an analytical tool to characterize nucleotide sequences. As such, for detection of mutations, LSSP-PCR combines in sensitivity and specificity the best features of shotgun strategies and techniques dependent on prior knowledge of base alterations and thus appears superior to all alternatives except DNA sequencing, although it has the advantages of being simpler and faster than the latter.

We found in preliminary experiments that LSSP-PCR was sensitive to several variables. In addition to the requirement for low annealing temperatures, very high concentrations of both the primer and *Taq* polymerase were essential for good results. It has been proposed that PCR is governed kinetically by two variables: the rate of primer dissociation from the primer–template complex before polymerization and the rate at which the DNA polymerase extends the primer until a stable duplex is formed (19). Our results are consistent with this, since the low annealing temperature and high concentrations of primer appear essential to drive the formation of primer–template complexes in conditions of considerable sequence mismatch, while high concentrations of polymerase are necessary to ensure a rapid extension reaction before dissociation occurs. Small sequence changes in the template would lead to an altered outcome in the competition of highly mismatched potential priming sites for available primers. This can explain why there are differences in many bands between the mutants even though they differ in a single-base substitution. An interesting feature of the signatures is several DNA fragments larger than the template. These possibly result from initial amplification products serving as primers in later rounds of amplification.

All our observations are consistent with the priming events that are responsible for the amplifications of the complex set of reaction products being very unstable. Due to this, one major concern was with the reproducibility of LSSP-PCR. However, as shown in Fig. 1, the signatures are reproducible. Nevertheless, in our routine use of the technique, we always try to use simultaneous amplification and concomitant electrophoretic analysis.

The striking and unexpected features of the simple LSSP-PCR protocol that we have developed for probing DNA sequences are the complexity of the banding patterns that are produced from the DNA fragments and the large variation that one or very few base changes can cause in the signatures. Alteration in a single base among 569 in the case of the myosin light chain gene, for example, resulted in almost half the bands being different. To account for the variations observed, we must conclude that the interactions under the very low stringency PCR conditions are very weak and unstable and that potential priming sites compete for available primers. This makes the amplification reaction exquisitely sensitive to alterations in template sequence, so that the test becomes an extremely powerful and rapid means of sequence comparison in fragments of up to the order of 1000 bp. This is clearly exemplified by our results with the D-loop region of human mtDNA, which showed matrilineage-specific signatures. Thus, LSSP-PCR may be used as a very useful screening test for genetic identity and establishment of matrilineal relationships, prior to the more definitive study of this region by DNA sequencing (16).

In conclusion, LSSP-PCR should be of value in essentially all situations where the rapid detection of variable and unpredictable changes in defined sequences is required. In all the cases we have so far investigated, LSSP-PCR has always produced complex patterns of fragments that are sensitive to sequence changes, and we find it reasonable to suppose that it will be universally applicable.

Part of this research was funded by grants-in-aid from the Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG).

1. Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1986) *Nature (London)* **322**, 90–93.
2. Saiki, R. K., Walsh, P. S., Levenson, C. H. & Erlich, H. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6230–6234.
3. Orita, M., Suzuki, Y., Sekiya, T. & Hyashi, K. (1989) *Genomics* **5**, 874–879.
4. Sheffield, V. C., Cox, V. C. & Myers, R. M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 232–236.
5. Stoneking, M., Hedgecock, D., Higushi, R. G., Vigilant, L. & Erlich, H. A. (1991) *Am. J. Hum. Genet.* **48**, 370–382.
6. Pena, S. D. J., Macedo, A. M., Gontijo, N. F., Medeiros, A. C. & Ribeiro, J. C. (1991) *Electrophoresis* **12**, 146–152.
7. Vigilant, L., Pennington, R., Harpending, H., Kocher, T. D. & Wilson, A. C. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9350–9354.
8. Santos, F. R., Pena, S. D. J. & Epplen, J. T. (1993) *Hum. Genet.* **90**, 655–656.
9. Reinach, F. C., Nagai, K. & Kendrick-Jones, J. (1986) *Nature (London)* **322**, 80–83.
10. Rosenthal, W., Seibold, A., Antaramian, A., Lonergan, M., Arthus, M.-F., Hendy, G. N., Birnbaumer, M. & Bichet, D. G. (1992) *Nature (London)* **359**, 233–235.
11. van den Ouweland, A. M. W., Dreesen, J. C. F. M., Verdijk, M., Knoers, N. V. A. M., Monnens, L. A. H., Rochhi, M. & Oost, B. A. (1992) *Nature Genet.* **2**, 99–102.
12. Pan, Y., Metzenberg, A., Das, S., Jing, B. & Gitschier, J. (1992) *Nature Genet.* **2**, 103–106.
13. Birnbaumer, M., Seibold, A., Gilbert, S., Ishido, M., Barberis, C., Antaramian, A., Brabet, P. & Rosenthal, W. (1992) *Nature (London)* **357**, 333–335.
14. De Marco, L., Bale, A. E., Carson, E., Boson, W., Nordenskjöld, M., Ritzén, M., Ferreira, P. C. P. & Friedman, E. (1993) *Am. J. Hum. Genet.* **53**, Suppl., 423.
15. Horai, S., Kondo, R., Murayama, K., Hayashi, S., Koike, H. & Nakai, N. (1991) *Philos. Trans. R. Soc. London B* **333**, 409–417.
16. Ginther, C., Issel-Tarver, L. & King, M.-C. (1992) *Nature Genet.* **2**, 135–138.
17. Welsh, J. & McClelland, M. (1990) *Nucleic Acids Res.* **18**, 7213–7218.
18. Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. & Tingey, S. V. (1990) *Nucleic Acids Res.* **18**, 6531–6535.
19. Wu, D. Y., Ugozzoli, L., Pal, B. J., Qian, J. & Wallace, R. B. (1991) *DNA Cell Biol.* **10**, 233–238.