GENOME METHODS

PCR Induction of a *TaqI* Restriction Site at Any CpG Dinucleotide Using Two Mismatched Primers (CpG-PCR)

Sandra D. O'Dell,¹ Steve E. Humphries, and Ian N.M. Day

Division of Cardiovascular Genetics, Department of Medicine, University College London Medical School, The Rayne Institute, London WC1E 6JJ, UK

A final common pathway has been devised for analysis of de novo mutation at any CpG site. Artificial restriction sites can be introduced in known DNA sequences by using either or both sense and antisense mismatched PCR primers. Siting of the primers directly adjacent 5' and 3' to the CpG site yields a 52-bp PCR product, resulting from the sum of the two 25-mer oligonucleotides plus the two intervening bases (C and G), and also yields consistent digestion fragments. Three out of four possible four-base palindromes (*TaqI*, *HhaI*, and *MspI*) were investigated for mutations R329X and E80K in the human LDL receptor gene, and for mutations R395W and R612C, the *TaqI* site was forced using PCR in which both primers had 3' mismatched T. Both empirically and on theoretical grounds, *TaqI* is the forced restriction site of choice. The approach has been adapted to the high-throughput microplate diagonal gel electrophoresis (MADGE) system for selective analysis of mutations at CpG sites, which may account for 20% of all single base variation in the human genome.

In combination with PCR amplification, restriction enzyme analysis forms a rapid and general technique to identify gene lesions that change a restriction enzyme site. In general, the combined PCR/restriction enzyme method is a two-step process. First, a set of specific primers is annealed to the target DNA to give a specific PCR product, that is, one single band in analytical gel electrophoresis. Second, after the PCR product has been digested with a known restriction enzyme and subjected to electrophoresis, a specific pattern of bands appears on the gel. These diagnostic criteria have been applied to a number of known mutations in single-copy genes (e.g. Saiki et al. 1985; Friedman et al. 1990). There are >100 different restriction enzyme recognition sequences known. Approximately one-sixth of the base pairs in the human haploid genome are covered by a naturally occurring restriction enzyme recognition site (Eiken et al. 1991). Less than 20% of the mutations would then be directly detectable by an available restriction enzyme, and polymorphisms or mutations without a restriction site were formerly only detected using allele-specific

radiolabeled probes. In such instances, appropriate restriction sites can be artificially created by site-directed mutagenesis in the amplification step (Haliassos et al. 1989), referred to as amplification created restriction sites (ACRS) (Eiken et al. 1991). The technique was first used to detect which specific mutations of *ras* oncogenes were present in tumor DNA (Haliassos et al. 1989; Jiang et al. 1989; Kumar and Dunn 1989). Eiken et al. (1991) estimate that the probability of finding a natural or creating an artificial restriction site encompassing a given single-base mutation and extending maximally three bases to each side is 98.4%. In general, there will be a choice of designed mismatch primers for use in diagnosis between any two alleles of known sequence. Davidow (1992) has devised a computer program to detect restriction sites that could be designed into the wild-type PCR product and not the mutant or, if need be, vice versa. The former represents a sensitive assay for de novo mutation screening (i.e., no false negatives), and the latter represents a specific assay for a known mutation (i.e., no false positives). The creation of an artificial restriction site, either in the wild-type or mutated gene, is being applied in an increasing number of cases (e.g., Haliassos et al. 1989; Ward et al. 1990; Eiken et al. 1991; Jacobson and

¹Corresponding author.
E-MAIL sodell@med.ucl.ac.uk; FAX + 44 71 209 6212.

Moskovits 1991; Ng et al. 1991; Berger et al. 1993; Hossle et al. 1994; Jardine et al. 1994; Rootwelt et al. 1994; Athma et al. 1995).

Novel specific cutting sites are generated by introducing a single nucleotide mismatch into the 3' end of an oligonucleotide primer that is itself positioned immediately adjacent to the mutation site. After incorporation of the primer into a PCR product, the variant nucleotide of one allele matches the sequence of the primer to introduce a restriction enzyme recognition site. The corresponding nucleotide of the other allele does not complement the primer to form a restriction site. After restriction enzyme digestion of the PCR product, cleavage will occur completely, partially, or not at all, and this identifies which alleles were present in the template DNA. The genetic diagnosis for presence or absence of the specific mutation can be made on the basis of simple restriction digest of the product followed by gel electrophoresis, rather than by the more cumbersome use of allele-specific oligonucleotide binding (Wallace et al. 1981), even though the chromosomal DNA sequences of the two alleles do not differ by the presence or absence of any known restriction site.

Here, we show how artificial restriction sites can be established in known DNA sequences by using either sense or antisense or both PCR primers to force a restriction site through mismatch of the 3'-terminal base with the target. In each case, forcing or not, both primers are sited directly adjacent to the target CpG site. This approach provides a possible general basis for identifying mutations (most commonly involving the C → T or G → A transition at the mutation-prone CpG site), when flanked by any combination of two bases one 5' and the other 3'. The approach was attempted first in three out of four possible combinations of four-base palindromes for R329X in exon 7 (Solberg et al. 1994) to create *TaqI*, *HhaI*, and *MspI* sites that together involved both single and double mismatched primers and was repeated at a second CpG mutation site in the LDL receptor gene, E80K (Webb et al. 1992). At both R329X and E80K, only a mismatched sense primer was required to force the *TaqI* site. To develop a general technique to identify mutations at CpG sites flanked by any combination of two bases, two further sites were investigated, R395W and R612C, (I.N.M. Day, unpubl.), in which a 3' mismatch was required on both sense and antisense primers positioned directly adjacent to the CpG dinucleotide to create a *TaqI* site.

UNIVERSAL ASSAY FOR MUTATIONS AT CpG SITES

RESULTS

Digestion of the PCR product yields two comigrating 25-bp fragments (with a 2-bp 5' overhang) in the normal sequence that contains the forced enzyme cutting site. This is, therefore, the pattern in normal homozygotes. In heterozygotes, only the PCR duplex consisting of two normal strands is cut. The two heteroduplexes and the homoduplex mutant strands do not contain a recognition site as a result of the mutation and so remain uncut. Samples from heterozygous individuals therefore appear as two bands, a 52-bp uncut fragment and two comigrating cut 25-bp fragments. (Homozygous mutants would appear as a single uncut band, but no FH homozygotes are present in the screened sample.)

R329X CpG Mutation Site

TaqI Site

The results of genotyping in four R329X heterozygotes and four normal homozygotes are shown in Figure 1a. *TaqI* digestion of the 52-bp product yielded two fragments of 25 bp (plus a 2-bp 5' overhang) in samples where a clear PCR product band was obtained in the checking gel. The comigrating 25-bp products migrate well ahead of the 52-bp uncut band. Normal homozygotes appear as a single 25-bp band and heterozygotes as two bands of 52 and 25 bp, providing unequivocal identification of genotype.

HhaI Site

The results of genotyping in four R329X heterozygotes and four normal homozygotes are shown in Figure 1b. The 52-bp product was expected to yield two fragments of 25 bp (plus a 2-bp 5' overhang) on digestion. On checking the efficiency of the PCR, all samples showed a single product of ~50 bp that was undigested by the enzyme. Furthermore, the identical product appeared in negative controls without template DNA. The same result was obtained over a range of MgCl₂ concentrations between 1.0 and 2.5 mM in the PCR.

MspI Site

The results of genotyping in four R329X heterozygotes and four normal homozygotes are shown

Figure 1 Polyacrylamide gel electrophoresis of PCR-amplified and restriction endonuclease-digested fragments: (a) forced *TaqI* site, (b) forced *HhaI* site, and forced *MspI* sites at the R329X site of the LDL receptor gene. An ethidium bromide-stained 7.5% polyacrylamide MADGE gel (see Methods) was used to analyze four heterozygous mutant (rows A–D) and four homozygous normal individuals (rows E–H). (a) *TaqI* digestion of the 52-bp product yielded two comigrating fragments of 25 bp (plus a 2-bp 5' overhang). Heterozygous mutants appear as two bands of 52 bp and 25 bp (A1–D1), normal homozygotes appear as a single 25-bp band (E1–H1), and undigested PCR products are shown in column 2. (b) At the forced *MspI* site, indistinct products of 47–61 bp were obtained in samples (heterozygous mutants A1–D1; normal homozygotes E1–H1) and negative control (A5) and were also resistant to enzyme digestion (column 2). At the forced *HhaI* site a single product of 47 bp (sized on a sequencing gel) was obtained in samples (heterozygous mutants A3–D3; normal homozygotes E3–H3) and in negative control (B5) and were resistant to enzyme digestion (column 4). Size markers of 65 bp and 84 bp are shown in C5, D5, and E5.

in Figure 1b. The 52-bp product produced with both priming oligonucleotides (FH 199 and FH 200) forcing mismatches at their 3' termini to create the restriction site was expected to yield two fragments of 25 bp (plus a 2-bp 5' overhang) on digestion. On checking the efficiency of the PCR, a single but diffuse band appeared of ~50 bp and, as with amplification at the *HhaI* site, appeared in negative controls without template DNA and was not digested by the enzyme. The same result was obtained over a range of $MgCl_2$ concentrations between 1.0 and 2.5 mM in the PCR.

E80K CpG Mutation Site

TaqI Site

In four E80K heterozygotes and four normal ho-

mozygotes *TaqI* digestion of the 69-bp product yielded fragments of 30 and 39 bp. Normal homozygotes appear as two cut 30- and 39-bp bands, and heterozygotes as 69-, 30-, and 39-bp bands (results not shown), providing identification of genotype with the same reliability as at the R329X site.

HhaI Site

In four E80K heterozygotes and four normal homozygotes, the 52-bp product produced with both priming oligos (FH 222 and FH 225) forcing mismatches at their 3' termini to create the restriction site was expected to yield two 25-bp fragments (plus a 2-bp 5' overhang) on digestion of the normal strand. On checking the efficiency of the PCR, strong bands of ~50 bp were found with normal and heterozygous DNA templates

and in negative controls. As at the R329X site, all were undigested by the enzyme, suggesting the presence of nonspecific, misprimed products (results not shown).

It therefore seemed likely that the products obtained at both R329X and E80K mutation sites when forcing a *HhaI* site were nonspecific and, coincidentally, appeared to be similar in size to the target amplified sequence. The product could have resulted from mispriming by either oligonucleotide, or primer-dimer formation. To investigate these possibilities, the precise size of the nonspecific product was measured by running [32 P]dCTP-labeled PCR single strands from negative control and normal DNA and heterozygous DNA templates alongside a [35 S]ATP-labeled M13 DNA sequence on a denaturing gel. The results showed a product of 47 bp was produced in PCRs with normal and heterozygous DNA and in negative controls.

MspI Site

In four E80K heterozygotes and four normal homozygotes, the 52-bp product produced with both priming oligos (FH 224 and FH 223) forcing mismatches at their 3' termini to create the restriction site was expected to yield two 25-bp fragments (plus a 2-bp 5' overhang) on digestion of the normal strand. On checking the efficiency of the PCR, bands of ~50 bp were seen in all heterozygotes and in normal samples. A number of bands in the 50-bp region were seen in the negative control. All products proved to be refractory to enzyme digestion (results not shown).

To investigate the possibility of mispriming or primer-dimer formation, the precise size of the nonspecific product was measured as previously by running [32 P]dCTP-labeled PCR single strands from negative control and normal DNA and heterozygous DNA templates alongside M13 samples on a sequencing gel. The results showed that a range of products between 61 and 47 bp were formed with normal and heterozygous DNA and negative controls.

R395W and R612C CpG Mutation Sites

TaqI Site

The results of genotyping in four R395W heterozygotes, one R612C heterozygote (tested in quadruplicate), and four normal homozygotes are

UNIVERSAL ASSAY FOR MUTATIONS AT CpG SITES

shown in Figure 2. *TaqI* digestion of the 52-bp product yielded comigrating fragments of 25 bp (plus a 2-bp 5' overhang). Normal homozygotes appear as a single 25-bp band, and heterozygotes as two bands of 52 bp and 25 bp. Hence, unequivocal genotype identification was demonstrated at two sites in which the forced *TaqI* site resulted from amplification involving two mismatched primers.

DISCUSSION

Artificial restriction fragment length polymor-

Figure 2 Polyacrylamide gel electrophoresis of PCR-amplified and *TaqI*-digested fragments at the R395W site and R612C site of the LDL receptor gene. An ethidium bromide-stained 7.5% polyacrylamide MADGE gel was used to analyze four R395W heterozygous mutants and one R612C heterozygous mutant in quadruplicate (rows A–D) and four homozygous normal individuals (rows E–H). R395W heterozygous mutants appear as two bands of 52 bp and 25 bp (B1–D1; A1 is failed amplification), normal homozygotes appear as a single band of two comigrating 25-bp fragments (E1–H1) and undigested PCR products are shown in column 2. At the R612C site, the same pattern occurs as at R395W. Heterozygous mutants are in A3–D3, normal homozygotes are in E3–H3, and undigested PCR products are shown in column 4.

phisms (RFLPs) are generated by introducing a single nucleotide mismatch into the 3' end of an oligonucleotide primer that is itself positioned immediately adjacent to the mutation site. After incorporation of the primer into a PCR product, the variant nucleotide of one allele matches the sequence of the primer to introduce a restriction enzyme recognition site, whereas the corresponding nucleotide of the other allele does not complement the primer and the site is absent. A limitation of the method is that a restriction site must be created within the primer in the confines of the genomic DNA sequence; for some mutations there may not be a potential restriction site to engineer. However, we have shown that by use of 3'-mismatched bases on both sense and antisense primers against the target (R395W and R612C), any combination of bases flanking the CpG site can be substituted by T upstream and A downstream of the dinucleotide, to create a T/CGA *TaqI* recognition site. The relative success we have achieved in forcing a *TaqI* site in comparison with the alternative *HhaI* or *MspI* site is discussed below.

Efficiency of Oligonucleotide Priming and Extension

The ability of an oligonucleotide to serve as a primer for PCR is dependent on several factors, including (1) the kinetics of association and dissociation of primer–template duplexes at the annealing and extension temperatures, (2) the effects on duplex stability of mismatched bases and their location, and (3) the efficiency with which the polymerase can recognize and extend a mismatched duplex.

Several investigators have evaluated the effect of single 3'-terminal mismatches on the efficiency of PCR (Ehlen and Dubeau 1989; Newton et al. 1989; Nichols et al. 1989; Okayama et al. 1989; Wu et al. 1989; Kwok et al. 1990). Because mismatches at or near the terminal 3' base of a primer are known to affect both oligonucleotide annealing stability and efficiency of polymerase extension, they would be expected to affect PCR more dramatically than mismatches at other positions (Petruska et al. 1988). Our successful forcing of a *TaqI* site, (T/CGA) critically depends on using a T in the 3' mismatch position on both sense and antisense primers adjacent to the CpG. Significantly, Kwok et al. (1990) found mismatches of T with either G, C, or T had minimal effect on PCR product yield; however, 3' primer

G mismatched to target was found to be among the most problematic (Petruska et al. 1988; Kwok et al. 1990). It appears that our unsuccessful attempts at forcing *HhaI* and *MspI* sites are attributable to the involvement of mismatched 3' G and C bases, respectively. Forcing a *HhaI* site (G/CGC) always involves a 3'-terminal G on both sense and antisense primers, and forcing an *MspI* site (C/CGG) always involves a 3'-terminal C on both sense and antisense primers. The relevance of mismatched G and C to the failure of the *HhaI* and *MspI* forced PCRs seems to relate to two factors. First, the likelihood of mispriming events with undesired complementary bases in the other primer, rather than the desired target mismatched bases, is encouraged by the formation of three hydrogen bonds between a 3'-terminal G and target C or the reciprocal 3' C and target G. The 3'-terminal T of the forced *TaqI* site would be less susceptible to mispriming events, as only two hydrogen bonds are involved in the T/A pairing. Second, according to Kwok et al. (1990) and Petruska et al. (1988), mismatches involving primer G and C exhibit varying degrees of efficiency in annealing and extension during PCR. The absence of misprimed products in negative controls at forced *TaqI* sites suggests that specific priming is largely unaffected when the 3' primer mismatch involves a T residue. The presence of misprimed products in the forced *HhaI* and *MspI* site PCRs seems to reinforce the finding that mismatches involving, respectively, G and C at the 3' terminus of the primer are less stable (Gaffney and Jones 1989) and are more likely to lead to mispriming.

Primer–dimer Formation

According to Kwok et al. (1990), a major product that migrates at approximately the distance expected of the sum of the two primers is often observed in amplification reactions, termed "primer–dimer." As the positioning of the primers adjacent to the mutation site is obligatory in forcing a recognition site, there is no possibility of resiting primers to avoid primer–dimer formation resulting from complementary sequences between the two primers. However, the possibility exists of altering a base remote from the sensitive 3' terminus to abolish complementation without diminution of PCR amplification efficiency. According to Kwok et al. (1990), mismatches 1, 2, or 3 bases from the 3' nucleotide of primers had no apparent effect on PCR product

yield. It appears likely that in the case of the forced *HhaI* PCR at the R329X site, the three 3'-terminal bases GCT of the 25-bp antisense primer (FH 189) may have paired with the first three 5' bases CGA of the 25-bp sense primer (FH 201) in forming a 47-bp primer-dimer product. Altering a base remote from the recognition site would not affect propensity to dimerize in this case. There does not appear to be the same opportunity for terminal base-pairing in the case of the forced *MspI* PCR primers at the same site, yet a range of nonspecific products between 61 and 47 bp were formed. These are more likely to reflect mispriming events.

Obligatory siting of the primers also carries a risk of incorporating a naturally occurring *TaqI* site in the amplified product. The risk is small, because an amplicon of only 52 bp is generated. However, for reasons described above, the efficiency of the amplification should not be compromised if a single base change were to be made in a primer to destroy any natural site, cleavage at which would result in fragments too small to detect.

In conclusion, the reliability of the assay forcing a restriction site by 3' modification of PCR primers, first, lies in the generation of a specific amplification product that can be checked for correct fragment length, and, second, the restriction enzyme analysis of the amplification product generates DNA fragments whose appropriate sizes can be examined by gel electrophoresis. The versatility of this new approach lies in siting both primers directly adjacent to the CpG site thereby reducing the post-PCR analysis to a common pathway, namely a *TaqI* digest and electrophoretic analysis of products that are of invariant size. The approach is also compatible with high-volume genotyping or screening programs, using utilities that we have described previously for simplifying PCR or restriction digest set-up (Bolla et al. 1995; Day et al. 1995; Whittall et al. 1995) and for achieving very high throughput polyacrylamide gel electrophoresis (Day and Humphries 1994; Bolla et al. 1995; O'Dell et al. 1995). Owing to variable kinetics of mismatches, it appears that, for reasons discussed above, those involving a 3' T on the forcing primer are likely to be more successful in amplification of target sequences than those involving C or G. The fourth possibility, a 3' mismatched A on both primers would generate the sequence ACGT at a CpG site, which is the recognition sequence of *MaeII*; however, this option was not considered

UNIVERSAL ASSAY FOR MUTATIONS AT CpG SITES

as the enzyme is extremely expensive. From the viewpoint of a generalized approach to the detection of point mutations at vulnerable CpG sites, the relative efficiency of the 3' T mismatch therefore indicates that *TaqI* is the forced restriction site of choice. It should be noted that the technique detects variation at either of the bases C or G and thus is sensitive to any CpG site mutation; the precise change would need to be established by sequencing.

METHODS

Oligonucleotides

Oligonucleotides were from Genosys, Cambridge, UK. The bases in heavy type indicate those forcing a restriction site in the amplified product by mismatch at the 3' terminus.

Forcing *TaqI* Recognition Site

The PCR primers used to amplify a 52-bp sequence encompassing the R329X site were FH 188 (sense primer), 5'-CGACGGCTTCCAGCTGGTGGCCCAT-3' and FH 203 (antisense primer), 5'-CACCGGGGAATCACCTTCGCATCTT-3'. The PCR primers used to amplify a 69-bp sequence encompassing the E80K site were those used by P. Wenham (unpubl.): FH 198 (sense primer), 5'-GGCCAAGTGGACTGCGACAACGGCTCAGT-3' and FH 23 (antisense primer), 5'-ATAGCAAAGGCAGGGCCACCTTAC-3'. The PCR primers used to amplify a 52-bp sequence encompassing the R395W site were FH 232 (sense primer), 5'-CGAGGTCAGGAAGATGACGCTGGAT-3' and FH 233 (antisense primer), 5'-GGGGATGAGGCTGTGTACTCGCTT-3'. The PCR primers used to amplify a 52-bp sequence encompassing the R612C site were FH 234, 5'-CAACGAAGCCATTTCAGTGCCAAAT-3' and FH 235, 5'-CAAGTTGACATCGGAACCTGTGAGT-3'.

Forcing *HhaI* Recognition Site

The PCR primers used to amplify a 52-bp sequence encompassing the R329X site were FH 201 (sense primer), 5'-CGACGGCTTCCAGCTGGTGGCCAG-3' and FH 189 (antisense primer), 5'-CACCGGGGAATCACCTTCGCATCTG-3'. The PCR primers used to amplify a 52-bp sequence encompassing the E80K site were FH 222 (sense primer), 5'-AAGTGGACTGCGACAACGGCTCAGG-3' and FH 225 (antisense primer), 5'-GGGCCACACTTACGACAGCCTTGCG-3'.

Forcing *MspI* Recognition Site

The PCR primers used to amplify a 52-bp sequence encompassing the R329X site were FH 199 (sense primer), 5'-CGACGGCTTCCAGCTGGTGGCCCA-3' and FH 200 (antisense primer), 5'-CACCGGGGAATCACCTTCGCATCTC-3'. The PCR primers used to amplify a 52-bp se-

quence encompassing the E80K site were FH 224 (sense primer), 5'-AAGTGGACTGCGACAACGGCTCAGC-3' and FH 223 (antisense primer), 5'-GGGCCACACTTACGACAGCCTTGCC-3'.

PCR

The positions of the primers, mismatched 3' bases, and restriction enzyme recognition sites at the different codons in the LDL receptor gene are shown in Figure 3. A summary of the mismatched 3'-terminal bases and outcome of the PCRs is given in Table 1.

Forcing *TaqI* Recognition Site

The sense oligonucleotides used to amplify the 52-bp sequence encompassing the R329X site or the 69-bp sequence at the E80K site involve a mismatched T at the 3' terminus of the primer. At R395W and R612C, both sense and antisense oligonucleotides similarly involve a 3'-terminal mismatched T. In so doing, the next three inserted bases C,G,A in the normal sequence create a recognition site T/CGA for the restriction endonuclease *TaqI* in the extension product. In each case, the mutation destroys the *TaqI* site.

Forcing *HhaI* Recognition Site

Both sense and antisense oligonucleotides used to amplify the 52-bp sequence encompassing the E80K site involve a mismatched G at the 3' terminus of each primer, as does the single antisense oligonucleotide required at the R329X site. In so doing, the two inserted bases C and G in the normal sequence create the *HhaI* recognition site GCG/C in the extension product. In each case, the mutation destroys the *HhaI* site.

Forcing *MspI* Recognition Site

At both the R329X and E80K sites, two mismatched oligonucleotides are required to create an *MspI* site. At both mutation sites, a mismatched C occurs at the 3' terminus of both sense and antisense primers in forming the 52-bp PCR product. In so doing, the two inserted bases C and G in the normal sequence create the *MspI* recognition site C/CGG. In each case, the mutation destroys the *MspI* site.

All PCR products were prepared using 20- μ l reactions with 20 μ l of oil in Omniplates on a Hybaid Omnigene, (Hybaid Ltd, Teddington, UK). PCR conditions for amplification of sequences encompassing the R329X and E80K sites were as described by Hobbs et al. (1992) or by the conditions below. For amplification of sequences encompassing the R395W and R612C sites, the following conditions were used: 1 cycle at 94°C for 5 min, 35 cycles at 55°C

for 1 min, 70°C for 1 min, and 94°C for 1 min; 1 cycle at 70°C for 1 min. The final short extension cycle was sufficient to complete amplification of the 52-bp product. Each 20- μ l reaction mixture contained 1.5 mM MgCl₂, 8 pmoles each PCR primer, estimated average 40 ng of human DNA, 0.2 units of *Taq* polymerase (GIBCO-BRL), 200 μ M each dNTP, 1% (vol/vol) DMSO (BDH, Poole, UK), 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 0.01% (wt/vol) gelatin. It should be noted that amplification with fidelity at the mutated base was unsuccessful using 0.05% w/v (GIBCO-BRL) in the PCR mixture both for R329X and E80K PCRs. In contrast to w/v, the use of DMSO permits fidelity; however, the omission of both DMSO and w/v is also successful.

To size PCR products accurately, some PCRs were labeled using 0.3 μ Ci of [α -³²P]dCTP (Amersham, UK) per 20 μ l PCR, prior to electrophoresis on sequencing gels.

Restriction Endonuclease Digestion of PCR Products

Restriction enzymes *TaqI*, *HhaI*, and *MspI*, BSA, and buffers were obtained from New England Biolabs (Beverly, MA). All enzymes were supplied at a concentration of 10 U/ μ l and diluted to one-tenth of this concentration in the digestion; that is, each 10- μ l reaction contained 1 unit of enzyme. The buffer was supplied at a 10 $\times$ concentration and diluted one-tenth in the digestion volume. The *TaqI* and *HhaI* digestion mixtures were supplemented with BSA diluted one-hundredth on addition to the digestion mixture to a final concentration of 100 μ g/ml. Each 10- μ l digestion was set up in an Omniplate. The *TaqI* and *HhaI* enzymes (10 U/ μ l), BSA (100 $\times$ diluted 10-fold in water to 1 mg/ml), and 10 $\times$ buffer were premixed in equal volumes, and 3 μ l of the mixture was added to the wells. PCR product (7 μ l) was then transferred by multichannel pipette and agitated thoroughly to mix. The *MspI* reactions were not supplemented with BSA, and 2 μ l of reaction mixture was mixed with 8 μ l of PCR product. The digestion volumes were covered with 10 μ l of paraffin oil, and the plate was covered with plastic film. Incubation of *TaqI* reactions was for 1 hr at 65°C on a Hybaid Omnigene heating block and of *HhaI* and *MspI* was for 1 hr at 37°C in an incubator.

Polyacrylamide Gel Electrophoresis of PCR and Digestion Products

Samples of undigested and digested PCR products were electrophoresed in adjacent wells, so that the efficiency of both PCR and digestion could be checked conveniently. A 7.5% polyacrylamide microplate diagonal gel electrophoresis (MADGE) gel (Day and Humphries 1994) was used to detect PCR products of 52 bp and digestion products of two 25-bp double-stranded fragments (with a 2-bp 5' overhang resulting from the staggered cutting site). Five microliters of PCR product or digestion product was mixed with 2 μ l of formamide loading buffer (Sambrook et al.

Figure 3 Amplified normal sequences, PCR primers, and restriction endonuclease enzyme digestion products at LDL receptor gene R329X, E80K, R395W, and R612C sites. In the normal sequences, the codon containing the mutation is underlined and the mutating base is indicated in boldface type. In the primers, the 3'-terminal base forcing a mismatch is indicated in boldface type. The forced enzyme recognition sites are indicated.

UNIVERSAL ASSAY FOR MUTATIONS AT CpG SITES

Figure 3 (See facing page for legend.)

Table 1. PCR Primer 3'-terminal Mismatches Forcing Different Restriction Sites at a Range of LDL Receptor Codons

Forced restriction site at LDL receptor codon	PCR primer 3'-terminal mismatch(es) (primer/template)	Observed outcome
<i>TaqI</i> at R329 codon	FH 188 sense T/C	faithful amplification, 52 bp
<i>HhaI</i> at R329 codon	FH 189 antisense G/A	misproduct, 47 bp
<i>MspI</i> at R329 codon	FH 199 sense C/C FH 200 antisense C/A	misproducts, 47–61 bp
<i>TaqI</i> at E80 codon	FH 198 sense T/T	faithful amplification, 69 bp
<i>HhaI</i> at E80 codon	FH 222 sense G/T FH 225 antisense G/A	misproduct, 47 bp
<i>MspI</i> at E80 codon	FH 224 sense C/T FH 223 sense C/A	misproducts, 47–61 bp
<i>TaqI</i> at R395 codon	FH 232 sense T/G FH 233 antisense T/G	faithful amplification, 52 bp
<i>TaqI</i> at R612 codon	FH 234 sense T/G FH 235 antisense T/C	faithful amplification, 52 bp

1989), and 5 μ l of this mixture was loaded onto the gel. Electrophoresis was at 150 V (10 V/cm) in TBE buffer for 30 min. It was found that digestion only proceeded satisfactorily if a clear single PCR product band was obtained. In these trials, eight samples consisting of four heterozygotes and four normal homozygotes were used for a given PCR. However, the unique format of the MADGE system allows convenient rapid checking of PCR products (including negative and positive controls) transferred directly from a standard 96-well Omniplate, on a single gel. A second gel would then be used to size the digestion products using the same electrophoresis conditions. Any equivocal digests can be cross-referenced with the PCR gel. After electrophoresis, the gel was viewed under UV transillumination.

Sizing of PCR Products

Two microliters of [α - 32 P]dCTP-labeled PCR product was mixed with 6 μ l of formamide dye and heated at 95°C on an Omnigene block for 3 min, to melt the duplex, and then placed on ice. Four microliters of the mixture was then loaded on a standard sequencing gel (Sambrook et al. 1989), alongside four tracks containing [35 S]ATP-labeled dideoxy-terminated fragments of a M13 sequence included in a Thermosequenase cycle sequencing kit (Amersham, UK). After electrophoresis, the gel was transferred to Whatman 3-mm chromatography paper, dried, and exposed to Betamax hyperfilm (Kodak) for 24 hr at -70°C . The precise length of the product could then be determined after autoradiography by identification of a comigrating band from a sequencing track.

ACKNOWLEDGMENTS

Dr. P. Wenham is thanked for disclosure prior to publication of conditions successful for generation of a *TaqI* site at

E80K using one mismatched oligonucleotide. S.D.O. is funded by Medical Research Council Project Grant G9432437MB awarded to I.N.M.D. and S.E.H. S.E.H. is funded by British Heart Foundation Chair and Programme grant PG95007. I.N.M.D. is a Lister Institute Research Fellow.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

- Athma, P., N. Fidahusein, and M. Swift. 1995. Single base polymorphism linked to the ataxia-telangiectasia locus is detected by mismatch PCR. *Biochem. Biophys. Res. Commun.* **210**: 982–986.
- Berger, J., B. Molzer, V. Gieselmann, and H. Bernheimer. 1993. Simultaneous detection of the two most frequent metachromatic leukodystrophy mutations. *Hum. Genet.* **92**: 421–423.
- Bolla, M., L. Haddad, A.F. Winder, S.E. Humphries, and I.N.M. Day. 1995. A method for the rapid determination of hundreds of APOE genotypes utilising highly simplified, optimised protocols and restriction enzyme digestion analysis by microplate array diagonal gel electrophoresis (MADGE). *Clin. Chem.* **41**: 1599–1604.
- Davidow, L.S. 1992. Selecting PCR designed mismatch primers to create diagnostic restriction sites. *Comput. Appl. Biosci.* **8**: 193–194.
- Day, I.N.M. and S.E. Humphries. 1994. Electrophoresis for genotyping: Microtitre array diagonal gel electrophoresis (MADGE) on horizontal polyacrylamide

UNIVERSAL ASSAY FOR MUTATIONS AT CpG SITES

(H-PAGE) gels, Hydrolink or agarose. *Anal. Biochem.* **222**: 389–395.

Day, I.N.M., R. Whittall, V. Gudnason, and S.E. Humphries. 1995. Dried template DNA, dried PCR oligonucleotides and mailing in 96-well plates: LDL receptor gene mutation screening. *BioTechniques* **18**: 981–984.

Ehlen, T. and L. Dubeau. 1989. Detection of ras point mutations by polymerase chain reaction using mutation-specific inosine-containing oligonucleotide primers. *Biochem. Biophys. Res. Commun.* **160**: 441–447.

Eiken, H.G., E. Odland, H. Boman, L. Skjelkvale, L.F. Engebretsen, and J. Apold. 1991. Application of natural and amplification created restriction sites for the diagnosis of PKU mutations. *Nucleic Acids Res.* **19**: 1427–1430.

Friedman, K.J., W.E. Highsmith, T.W. Prior, T.R. Perry, and L.M. Silverman. 1990. Cystic fibrosis deletion mutation detected by PCR-mediated site-directed mutagenesis. *Clin. Chem.* **36**: 695–696.

Gaffney, B.L. and R.A. Jones. 1989. Thermodynamic comparison of the base pairs formed by the carcinogenic lesion 06-methylguanine with reference both to Watson-Crick pairs and to mismatched pairs. *Biochemistry* **28**: 5881–5889.

Haliassos, A., J.C. Chomel, L. Tesson, M. Baudis, J. Kruh, J.C. Kaplan, and A. Kitzi. 1989. Modification of enzymatically amplified DNA for the detection of point mutations. *Nucleic Acid Res.* **17**: 3606.

Hobbs, H., M.S. Brown, and J.L. Goldstein. 1992. Molecular genetics of the LDL receptor gene in familial hypercholesterolaemia. *Hum. Mutat.* **1**: 445–466.

Hossle, J.P., M. de Boer, R.A. Seger, and D. Roos. 1994. Identification of allele-specific p22-phox mutations in a compound heterozygous patient with chronic granulomatous disease by mismatch PCR and restriction enzyme analysis. *Human Genet.* **93**: 437–442.

Jacobson, D.R. and T. Moskovits. 1991. Rapid, non-radioactive screening for activating ras oncogene mutations using PCR-primer introduced restriction analysis. *PCR Methods Applic.* **1**: 146–148.

Jardine, P.E., P.D. Cotter, S.A. Johnson, E.J. Fitzsimons, L. Tyfield, P.W. Lunt, and D.F. Bishop. 1994. Pyridoxine-refractory congenital sideroblastic anaemia with evidence for autosomal inheritance: Exclusion of linkage to ALAS2 at Xp11.21 by polymorphism analysis. *J. Med. Genet.* **31**: 213–218.

Jiang, W., S.M. Kahn, S.G. Guillem, S.H. Lu, and I.B. Weinstein. 1989. Rapid detection of ras oncogenes in human tumors: Applications to colon, esophageal and gastric cancer. *Oncogene* **4**: 923–928.

Kumar, N. and L. Dunn. 1989. Designed diagnostic

restriction fragment length polymorphisms for the detection of point mutations in ras oncogenes. *Oncogene Res.* **1**: 235–241.

Kwok, S., D.E. Kellog, N. McKinney, D. Spasic, L. Goda, C. Levenson, and J.J. Sninsky. 1990. Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. *Nucleic Acids Res.* **18**: 999–1005.

Newton, C.R., A. Graham, L.E. Heptinstall, S.J. Powell, C. Summers, N. Kalsheker, J.C. Smith, and A.F. Markham. 1989. Analysis of any point mutation in DNA. The amplification-refractory mutation system. *Nucleic Acids Res.* **17**: 2503–2516.

Ng, I.S.L., R. Pace, M.V. Richard, K. Kobayashi, K. Bat-sheva, L.-C. Tsui, and A.L. Beaudet. 1991. Methods for analysis of multiple cystic fibrosis mutations. *Hum. Genet.* **87**: 613–617.

Nichols, W.C., J.J. Liepnieks, V.A. McCusick, and M.D. Benson. 1989. Direct sequencing of the gene for Maryland/German familial amyloidotic polyneuropathy type II and genotyping by allele-specific enzymatic amplification. *Genomics* **5**: 535–540.

O'Dell, S.D., S.E. Humphries, and I.N.M. Day. 1995. A rapid approach to genotyping of the insertion/deletion polymorphism in intron 16 of the angiotensin converting enzyme using simplified DNA preparation and microtitre array diagonal gel electrophoresis. *Br. Heart J.* **73**: 368–371.

Okayama, H., D.T. Curiel, M.L. Brantly, M.D. Holmes, and R.G. Crystal. 1989. Rapid, non-radioactive detection of mutations in the human genome by allele-specific amplification. *J. Lab. Clin. Med.* **114**: 105–113.

Petruska, J., M. Goodman, M.S. Boosalis, L.C. Sowers, C. Cheong, and I. Tinoco. 1988. Comparison between DNA melting thermodynamics and DNA polymerase fidelity. *Proc. Nat. Acad. Sci.* **85**: 6252–6256.

Rootwelt, H., T. Kristensen, R. Berger, K. Hoie, and E.A. Kvittingden. 1994. Tyrosinemia type-1 complex splicing defects and a missense mutation in the fumarylacetoacetase gene. *Hum. Genet.* **94**: 235–239.

Saiki, R.K., S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, and H.A. Erlich. 1985. Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anaemia. *Science* **230**: 1350–1354.

Sambrook, J., E.F. Fritsh, and T. Maniatis. 1989. Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Solberg, K., O.K. Rodningen, S. Tonstad, L. Ose, and T.P. Lerer. 1994. Familial hypercholesterolaemia caused by a non-sense mutation in codon 329 of the LDL receptor gene. *Scand. J. Clin. Lab. Invest.* **54**: 605–609.

O'DELL ET AL.

Wallace, R.B., M.J. Johnson, T. Hirose, T. Miyake, E.H. Kawashima, and K. Itakura. 1981. The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequence to rabbit beta-globin DNA. *Nucleic Acids Res.* **9**: 879–894.

Ward, M., N.F. Olivieri, J. Ng, and J.C. Rode. 1990. Detection of β -thalassemia using an artificial-restriction fragment length polymorphism generated by the polymerase chain reaction. *Nucleic Acids Res.* **19**: 959.

Webb, J.C., X.-M. Sun, D.D. Patel, S.N. McCarthy, B.L. Knight, and A.K. Soutar. 1992. Characterization of two new point mutations in the low density lipoprotein receptor genes of an English patient with homozygous familial hypercholesterolaemia. *J. Lipid. Res.* **33**: 689–698.

Whittall, R., V. Gudnason, G. Weavind, L.B. Day, S.E. Humphries, and I.N.M. Day. 1995. Utilities for high throughput use of the single strand conformational polymorphism method: Screening of 791 patients with familial hypercholesterolemia patients for mutations in exon 3 of the low density lipoprotein receptor gene. *J. Med. Genet.* **32**: 509–515.

Wu, D.Y., L. Ugozzoli, B.K. Pal, and R.B. Wallace. 1989. Allele-specific enzymatic amplification of beta-globin genomic DNA for diagnosis of sickle-cell anemia. *Proc. Natl. Acad. Sci.* **86**: 2759–2762.

Received March 4, 1996; accepted in revised form April 24, 1996.

PCR induction of a TaqI restriction site at any CpG dinucleotide using two mismatched primers (CpG-PCR).

S D O'Dell, S E Humphries and I N Day

Genome Res. 1996 6: 558-568

Access the most recent version at doi:[10.1101/gr.6.6.558](https://doi.org/10.1101/gr.6.6.558)

References

This article cites 32 articles, 8 of which can be accessed free at:
<http://genome.cshlp.org/content/6/6/558.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Affordable, Accurate
Sequencing.

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>