Volume 16 Number 18 1988 Nucleic Acids Research

Recombinant fragment assay for gene targetting based on the polymerase chain reaction

Hyung-Suk Kim and Oliver Smithies

Laboratory of Medical Genetics and Genetics, University of Wisconsin, Madison, WI 53706, USA

Received May 31, 1988; Revised and Accepted August 18, 1988

ABSTRACT

The modification of chromosomal genes by homologous recombination between exogenous DNA and a target locus provides a powerful approach to the study of gene function. One of the current limitations of this gene targetting is the difficulty of identifying cells containing the correctly modified target locus when the modified gene is not amenable to either direct or indirect selection. We here describe a procedure for identifying correctly modified cells that depends on amplifying by the polymerase chain reaction (PCR) predictable fragments of DNA present only in the desired recombinants. This recombinant fragment assay can detect targetted modification using only a few cells, either alone or mixed with tens of thousands of unmodified cells; it does not depend on the phenotype of the modified gene, or on its expression in the target cells. The PCR amplification needed for the procedure is carried out with a heat stable polymerase and a simple automatic temperature-shift apparatus which is described.

INTRODUCTION

Homologous recombination between DNA introduced into recipient mammalian cells and a target chromosomal gene has been demonstrated in several laboratories (1-11). The potential usefulness of this gene targetting for modifying specific genes in predictable ways is obvious. However, because the desired recombinants typically occur at low frequencies, the targetting procedure has to date been largely confined to gene modifications leading to selectable phenotypes in cells in which the target locus is expressed. There is need for a procedure for detecting targetted modification that is not so constrained. One such procedure was described in the course of experiments that led to the targetted modification of a native human beta globin gene (6). In those experiments the desired recombinants were identified by looking for a specific recombinant fragment of DNA, present only in cells in which the planned recombination had occurred. A bacteriophage-mediated gene-rescue assay was employed to find the recombinant fragment, but the assay was laborious, and required the introduction of a prokaryotic gene into the target

© IRL Press Limited, Oxford, England.

Table 1. Oligonucleotides used as primers

Name[a]	Sequence	GC-content	Temperature[b]
Neo-left	(5')GTGTTCCGGCTGTCAGCGCA(3')	65%	70°C
Neo-right 1	(5')GTCCTGATAGCGGTCCGCCA(3')	65%	70°C
Neo-right 2	(5')CCATCGTGCCTCCCCACTCC(3')	70%	70°C
Neo-right 3	(5')CAGGAGGACACAGAGGGTGG(3')	65%	70°C
HPRT-left	(5')CAGCACAGGCAAGTAGAGAC(3')	55%	65°C
Neo-right 4	(5')TGCGCTGACAGCCGGAACAC(3')	65%	65°C

[a] The names refer to the text usage.
[b] The temperatures listed refer to the reaction temperature used with the indicated primer.

locus. In the present paper we describe experiments demonstrating that the same principle - the detection of a recombinant fragment of DNA - can be used in a much more general way to identify cells in which gene targetting has been successful. The new assay depends on the selective amplification of the recombinant DNA fragment by the polymerase chain reaction (PCR) (12,13). The assay does not require that the modified gene have a detectable phenotype, nor does it require one particular type of modification. Consequently the assay is applicable to a wide variety of genes in many cell types. A related scheme has been described for detecting chromosomal rearrangements (14).

MATERIALS AND METHODS

Nucleotide triphosphates as 100 mM aqueous solutions were from Pharmacia, Piscataway, NJ.

Oligonucleotide primers were synthesized by the Biotechnology Center, University of Wisconsin-Madison, on an Applied Biosystems DNA Synthesizer. They were normally used after deblocking but without further purification, except as indicated in the text. The dried deblocked nucleotides were taken up in 10mM Tris-HCl (pH8.0), 1mM Na_2EDTA and extracted once with buffer-saturated phenol/hydroxyquinoline (100:1), and once with chloroform/isoamyl alcohol (24:1). The sequences of six primers used in experiments reported here are given in Table 1, which also shows their GC-contents and the temperatures at which they were used.

Taq polymerase. Heat stable DNA polymerase from Thermus aquaticus was from several sources, including: New England Biolabs, Beverly, MA; Perkin Elmer

Cetus, Norwalk, CT; Stratagene, San Diego, CA; Epicentre Technologies, Madison, WI.

PCR mix (modified from the recommendations of New England Biolabs) is 67mM Tris-HCl (pH 8.8), 6.7mM $MgCl_2$, 16.6mM $(NH_4)_2SO_4$, 10mM 2-mercapto-ethanol, 6.7μM Na_2EDTA, 10% dimethylsulfoxide, 330 μM of each dNTP, 3.0 μg of each primer, 1 unit Taq polymerase per 50 μl.

Enzyme diluting buffer. 67mM Tris-HCl(pH8.8), 6.7mM $MgCl_2$, 16.6mM $(NH4)_2SO_4$, 10mM 2-mercaptoethanol, 6.7μM Na_2EDTA, 10%(v/v) glycerol.

Reaction protocol with cells. Take one to 10^5 cells in 5 μl of phosphate buffered saline. Add 20 μl distilled water. Incubate at 90°-95°C for 10 min. Centrifuge briefly to spin down condensate. Add 10 μg proteinase K (Boehringer Mannheim, Indianapolis, IN) in 1 μl distilled water. Incubate at 55°C for 30 min. Centrifuge briefly. Incubate at 90-95°C for 10 min. to destroy the proteinase K. Centrifuge briefly. Add 25 μl 2X concentrated PCR mix. Cover the sample with 20 μl light paraffin oil to stop evaporation. Amplify up to 30 cycles in the automatic temperature shift apparatus (Figure 2 below). In experiments with 50 cycles of amplification, 1 unit of Taq polymerase in 1 μl of enzyme diluting buffer is added after 25 cycles. At the end of the amplification, the sample is removed and partially dried under vacuum until the volume is about 20 μl. The samples are used directly for analysis by agarose gel electrophoresis.

Protocol with DNA solutions. Take up to 1 μg of genomic DNA or up to 1 ng of plasmid DNA in 1 μl of DNA dialysis buffer for use as template DNA. Add 24 μl of PCR mix. Add 20 μl light paraffin oil. Amplify.

Electrophoresis 2% electrophoresis-grade Agarose (Bethesda Research Laboratories, Gaithersburg, MD), 40mM Tris, 20mM NaCl, 20mM sodium acetate, 2mM Na_2EDTA, pH 8.0 with acetic acid. Electrophoresis is 1 to 2 hours at 5 V/cm.

RESULTS AND DISCUSSION

Principle of assay

The principle of the recombinant fragment assay is illustrated in Figure 1, in which we consider a simple crossover between two homologous DNA sequences giving rise to a recombinant. We imagine a target locus, represented by AB in the figure, being exposed to incoming DNA (CD) having nucleotide sequences in common with the target locus (hatched region). We are looking for cells in which homologous crossing over between these shared sequences has generated a recombinant (AD). We are interested in detecting the recombinant sequence AD

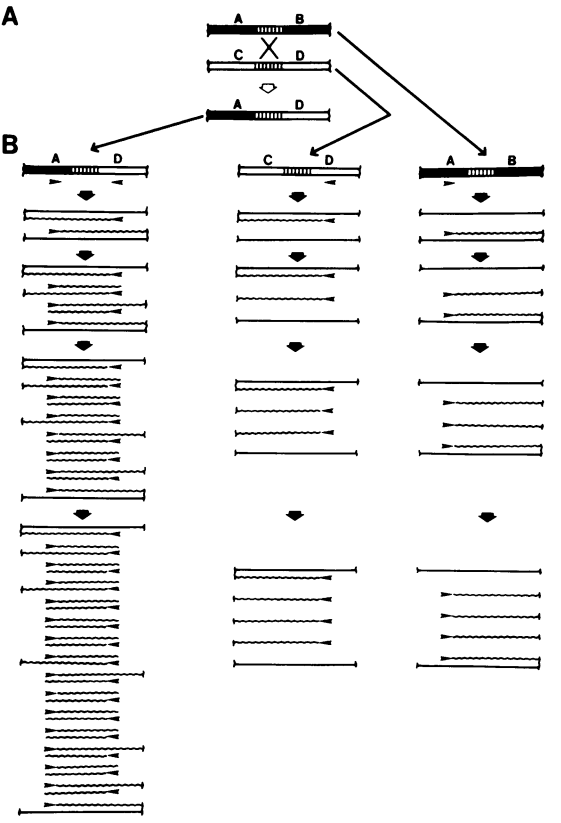

Figure 1. Recombinant fragment assay based on PCR amplification.
A. Homologous recombination between a target chromosomal sequence, AB (heavy black line), and exogenous DNA, CD (parallel lines), having nucleotide sequences in common with the target locus (hatched regions), to give a recombinant, AD.
B. The results of four successive rounds of PCR amplification in the presence of the two primers (indicated by the small arrowheads under the corresponding DNA sequences, A and D,) using template DNA from the recombinant chromosome, AD; the exogenous DNA, CD; and the target chromosome, AB. The wavy lines represent single strands of newly synthesized DNA. The straight lines represent single strands of the starting DNA. Each solid vertical arrow represents one round of amplification.

in the presence of an excess of the two "parental" sequences (AB and CD). (We discuss below, in considering Figures 4 and 7, how detection of a simple recombinant sequence allows the identification of targetting events of various categories.) To carry out the assay, DNA from cells or pools of cells in which recombinants are being sought is amplified by the PCR method in the presence of two primers - one specific for a part of each of the two parental

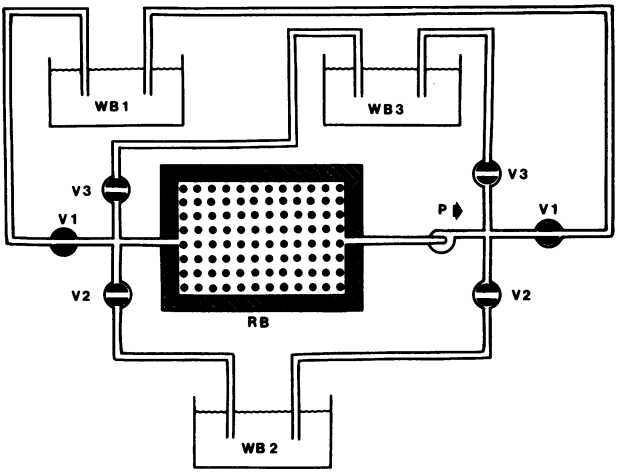

Figure 2. Schematic diagram of the automatic temperature-shift apparatus.
The temperature of the reaction block (RB) is controlled by water pumped by a centrifugal pump, P (Cole-Parmer, Model J-7004-54, Chicago, IL) from one of three water baths (WB1, WB2 and WB3). One pair of valves (V1) opens during the melt phase; a second pair (V2) opens during the anneal phase, and a third pair (V3) during the reaction phase. When a two-temperature protocol is used, the valves V3 are kept closed, and WB3 can be omitted. The valves (Automatic Switch Company, Model 8210C94, Florham Park, NJ) are controlled by a simple repetitive timer (circuit available on request) operating for a pre-set number of cycles with three independently adjustable consecutive ON/OFF time switches. The WB1 and WB3 baths (Blue M Company, Model MW-1120A-1, Blue Island, IL) each have a volume of about 35 liters; the lowest temperature bath, WB2, which must be capable of cooling as well as heating, has a volume of about 28 liters (Forma Scientific, Model 2095, Marietta, OH). The reaction block, RB, with tapering holes drilled on 0.5" centers to match the full length of 0.5 ml polypropylene Eppendorf tubes (Brinkman Instruments, Westbury, NY), is made of aluminum, external dimensions 5" x 8" x 1.75". Criss-crossing channels were machined inside the block to create a water-filled chamber that allows good heat transfer between the circulating water and the block. The inflow and outflow ports of the chamber have baffles to help distribute the water flow uniformly through the chamber. The reaction block is insulated on all sides with 7/8" rubber foam sheet.

DNA sequences participating in the recombination. The primer specific for A (part of the target locus) will allow synthesis of DNA sequences corresponding to the unmodified target locus, AB. Similarly the primer specific to D (part of the incoming DNA) will allow synthesis of DNA sequences corresponding to any of the incoming DNA, CD, that may either be present free in the treated cells or may have been inserted into their genomes at random positions. These syntheses, which utilize the parental sequences as templates, will be in linear proportion to the number of rounds of amplification, and the resulting

fragments will be single-stranded DNA fragments of indeterminate length, as illustrated in the figure. In contrast, synthesis of DNA corresponding to the correctly modified recombinant locus, AD, in which the two participating parental sequences are now colinear, will occur from both primers. The amplification will consequently be exponential, as illustrated in the figure, and will give rise to a double-stranded recombinant fragment of predictable length. After 20 rounds of amplification the predicted recombinant fragment will have been amplified up to a million fold (2^{20}), while the indeterminate parental sequences will have been amplified only twenty fold. Detection of the predicted recombinant fragment will establish that targetting has been achieved.

Temperature-Shift Apparatus

We have constructed a simple apparatus (Figure 2) for carrying out the PCR amplification. The apparatus repeatedly changes the temperature of up to 96 samples at once in a programmed way. This allows the DNA under test to be melted, the specific primers to be annealed to the resulting single-stranded DNA, and polymerase-catalyzed synthesis of DNA to occur from the primers. In order to avoid repeated additions of polymerase during the assay, we followed the suggestion of Mullis and Faloona (15) [see also (16) and (17)] and have used the heat stable polymerase, Taq polymerase (18, 19), from the thermophile Thermus aquaticus as the catalyst for the DNA synthesis.

Our apparatus was initially arranged to program three different temperatures, one each for the melt, anneal and reaction phases of the PCR method, but in our hands the results are indistinguishable when only two temperatures are used - one for the melting phase and the second for a combined annealing and reaction phase. Accordingly we present in Figure 2 a design amenable to either two or three temperatures. A related design has recently been described (20). Rapid shifting of the temperatures and keeping a short melt time help reduce exposure of the polymerase to high temperatures. The water circulation system and thermal masses of the system were therefore arranged (see legend to Figure 2) so that the melt phase can be executed rapidly in less than 60 seconds, of which only about 15 seconds are at the melt temperature of $90^{o} \pm 1^{o}C$. (The temperatures were measured by placing a small mercury thermometer surrounded by water in the holes of the reaction block.) We have not usually found it necessary to vary this time or temperature, although with some amplified fragments having regions with 70% GC-content we doubled the melt phase time. The anneal/reaction temperature, which is adjusted according to the GC-content of the primers (see below), is reached about 30 seconds after

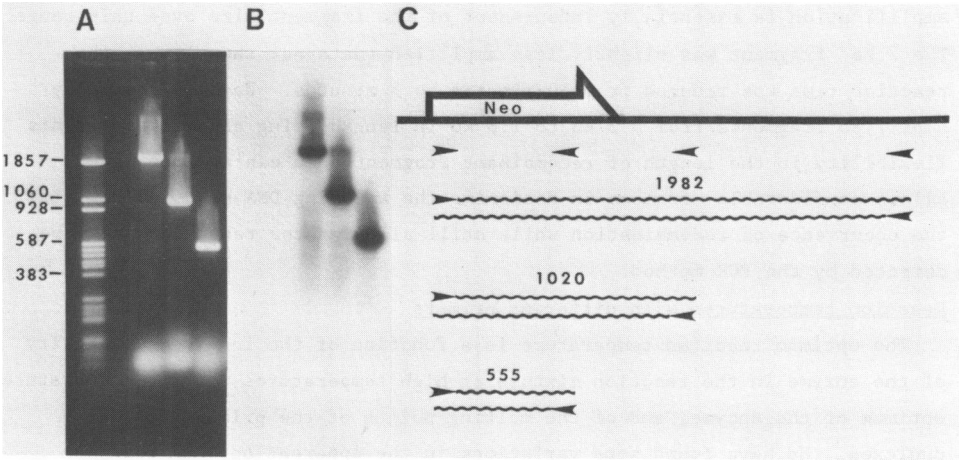

Figure 3. PCR amplification of fragments of different lengths.
A. Ethidium bromide-stained 2% agarose gel after electrophoresis of three PCR amplified fragments of DNA from pSV2Neo (21) using different pairs of primers. The sizes of some of the fragments in the marker mix (left most lane) are shown in base pairs.
B. Autoradiograph (10 min. exposure) of a Southern blot of the same gel after hybridization to a probe specific for an internal part of the Neo gene that does not include any primer sequences.
C. Diagramatic representation of the plasmid pSV2Neo showing the locations of the Neo gene (open large arrow) and of the three pairs of primers (small arrowheads) used in the three experiments. Table 1 in Materials and Methods gives the nucleotide sequences of the left primer (Neo-left) and of the three right primers (Neo-right 1, 2 and 3, from left to right). The expected sizes of the amplified fragments (parallel wavy lines) are given in base pairs.

completing the melt phase. Usual anneal/reaction temperatures are 60°C, 65°C or 70°C. The reaction time has been varied from 5 minutes, for fragments of 0.5 kb (kilobase pairs) length or less, to 15 minutes for fragments of several kb length.

Lengths of amplified fragments

In order to allow flexibility in the choice of primers for use in detecting recombination it is valuable to be able to amplify recombinant fragments of different lengths. We have therefore tested our ability to do this. Figure 3 illustrates the results obtained with fragments of 0.5 kb, 1 kb and 2 kb length. Thirty cycles of amplification were used with four different primers (three pairs), Neo-left, Neo-right 1, Neo-right 2 and Neo-right 3 (see Table 1). The template DNA was 50 pg of Eco RI-linearized pSV2Neo (21). The reaction time was 10 minutes per cycle at 70°C. The results show that the

amplification is essentially independent of the fragment size over this range. The 2 kb fragment was slightly less amplified (data not shown) when the reaction time was reduced from 10 minutes to 5 minutes. We have routinely amplified fragments from 0.3 kb to 1.8 kb in length using genomic DNA. This flexibility in the length of recombinant fragment that can be identified allows considerable freedom in designing the incoming DNA so as to maximize the occurrence of recombination while still allowing the recombinants to be detected by the PCR method.

Reaction temperatures with different primers

The optimum reaction temperature is a function of the long-term stability of the enzyme in the reaction mixture at high temperatures, of the temperature optimum of the enzyme, and of the melting points of the primer-template duplexes. We have found some variations in the apparent optimum reaction temperature with different preparations of the Taq polymerase. This variation is at least in part the consequence of different high temperature stabilities of the enzyme in different preparations. The GC-contents of the primers also influence the reaction temperature optimum. We have used 20 different 19 or 20 base oligonucleotides to prime the amplification of 16 different DNA fragments. The GC-contents of the primers have ranged from 26% to 70%. The optimum reaction temperature for each pair of primers was determined to the nearest 5°C by finding the temperature at which the desired fragment was maximally amplified without accompanying amplification of non-specific fragments. Occasionally yield had to be sacrificed to obtain specificity. All primer pairs (9 tested) with an average GC-content of less that 60% had an optimum temperature of 60°C, as did two pairs with 60% and 62.5% GC. Two primer pairs with GC-contents of 60% and 67.5% were optimal at 65°C. Two primer pairs with a GC-content of 65% and one pair with 67.5% GC were optimal at 70°C. Where possible we have tried to match the GC-contents of the two primers in any given amplification so that comparable melting points can be expected of each member of a pair. In this way at the optimum temperature both primers will be equal in their abilities to discriminate between their nominal targets and related but not identical sequences. We have avoided primers with obvious potential secondary structures. When a primer is being used to discriminate between two related sequences, we have found that the discrimination is better if the mismatches lie in the central region and/or towards the 3′ ends of the primers rather than in the 5′ region.

We tested primers that had been further purified, for example by HPLC, but

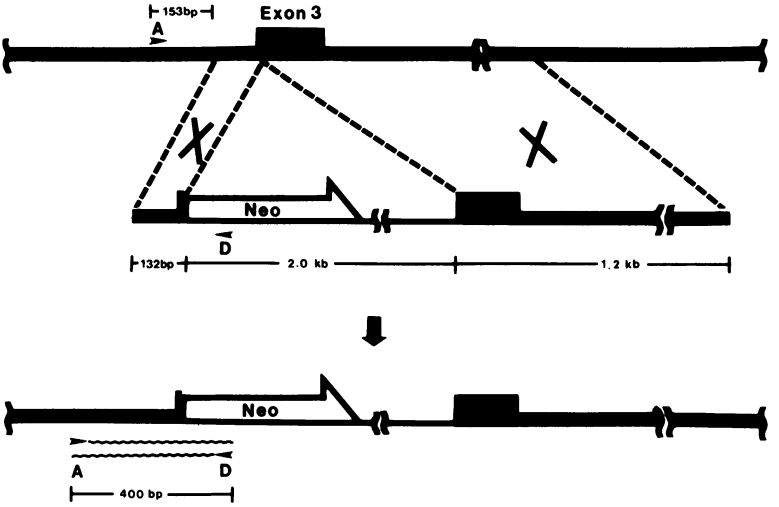

Figure 4. Use of the recombinant fragment assay to detect recombination at the HPRT-locus.

The top line depicts exon 3 and parts of introns 2 and 3 of the mouse native HPRT locus. Transcription is from left to right.

The middle line shows the incoming linear DNA used by Doetschman and his collaborators (22) to inactivate the HPRT gene by a targetted double crossover event. The broken lines high-light the 132 bp and 1.2 kb stretches of DNA that are identical in sequence in the target locus and incoming DNA. Cross-overs (X's) within the regions of identity lead to the recombinant.

The bottom line shows the recombinant and illustrates how it can be detected by the assay.

Heavy horizontal lines represent mouse genomic DNA sequences. The black rectangles show exon 3 or its 5' and 3' parts. The sizes of fragments of interest are indicated. The Neo gene (Neo) plus plasmid sequences (light horizontal lines) are not to scale. The arrowheads labelled A and D represent the primers (HPRT-left and Neo-right 4 (in Table 1) that allow amplification of the 400 bp recombinant fragment (parallel wavy lines, AD).

the results did not differ from those obtained with primers prepared as described under Materials and Methods.

Test System

As a reconstruction system for developing the recombinant fragment assay, we used modified cells in which targetted homologous recombination had already been successfully achieved in a selection experiment (22). The target locus chosen in the selection experiment was that coding for hypoxanthine phospho-ribosyl transferase (HPRT). The incoming DNA was a linear fragment so arranged that the native HPRT gene could be inactivated by a double crossover event between the target locus and the incoming DNA. Figure 4 depicts the

recombinational event, and illustrates how the PCR method can be used to detect the resulting recombinant. Two primers are used, one (A) from within intron 2 of the chromosomal HPRT target gene, and one (D) from within the Neo gene in the incoming DNA. The nucleotide sequences of these primers are given in Table 1 as HPRT-left and Neo-right 4. They allow efficient amplification of a specific 400 bp (base pair) fragment (AD) that is characteristic of the desired recombinants. Neither the unmodified chromosomal HPRT gene nor the incoming DNA, whether free in the treated cells or incorporated non-specifically into some irrelevant chromosomal locus, can generate the diagnostic 400 bp fragment. In what follows we refer to cells in which the HPRT gene has been modified as shown in Figure 4 as "HPRT-modified cells." Because the HPRT locus is on the X chromosome, and the modified cells are male-derived, each HPRT-modified cell carries a single copy of the modified gene.

Applicability of the assay to small numbers of cells

At various stages in the course of a gene targetting experiment with a non-selectable gene it is necessary to screen many small cultures or single colonies of cells. This screening can be greatly facilitated if it requires only a few cells, and a simple method for obtaining DNA from them. Elaborating on a procedure published by others (23), we therefore devised a simple water lysis/proteinase K digestion procedure (see Materials and Methods) that allows the direct use of cells in the recombinant fragment assay.

Figure 5 shows the results obtained when we tested the ability of our assay to detect the diagnostic HPRT recombinant fragment using a single HPRT-modified cell per reaction tube. The modified cells used were mouse embryonic stem cells that had been grown on feeder layers. Prior to the assay the embryonic stem cells were obtained in suspension by treating a culture with trypsin. Individual cells were observed microscopically and were taken up into a micropipette for introduction into separate reaction tubes. Amplification was for 50 cycles, and electrophoresis was carried out as described under Materials and Methods. Figure 5A shows that a visible band of the predicted size was detected by ethidium bromide staining in three tests with single cells. The authenticity of the stained band was confirmed by Southern blotting and hybridization to a probe specific for the HPRT gene (Figure 5B). In more extensive tests we have found that we can detect single recombinant cells in about 50% of the reaction tubes, and in all of the tubes with ten cells. Some of the negative results with single cells are probably

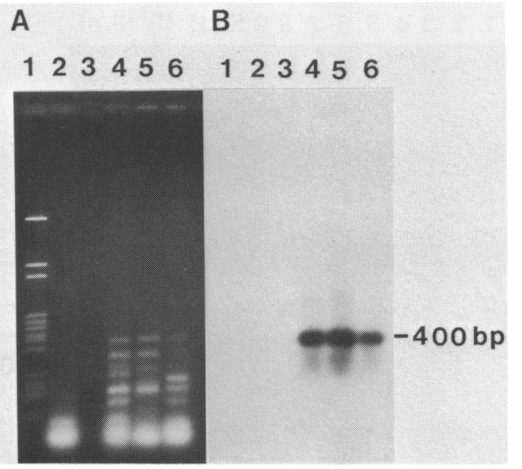

Figure 5. Detection of single recombinant cells.
A. Ethidium bromide stained 2% agarose gel after electrophoresis of three amplified samples each containing a single HPRT-modified cell (lanes 4, 5 and 6). Lane 1 is the size marker mix. Lane 2 is an amplified control sample (buffer control) without any added DNA. Lane 3 is empty.
B. Autoradiograph (overnight exposure) of a Southern blot of the same gel after hybridization to a fragment from the HPRT gene that does not include any primer sequences. The 400 bp recombinant fragment is indicated.

due to our inadvertently choosing a feeder layer cell, rather than an embryonic stem cell.

The extreme sensitivity of the procedure (in the present instance an ability to amplify only one template molecule to the level of a visible electrophoresis band) emphasizes the need for strict avoidance of contamination when working with the PCR method at its maximum sensitivity. Suitable controls without any added cells or with the addition of irrelevant cells (see below) should be included to check for possible contamination.

Of interest is the appearance of multiple non-specific bands in the reaction tubes containing single cells, in addition to the predicted fragment (lanes 4, 5 and 6, Figure 5A). We have frequently observed similar bands in such tests, and also in reaction tubes amplified without added DNA (the buffer controls) after extensive amplification. (See, for example, the buffer control in lane 1 of Figure 6, below.) The bands are usually spaced somewhat regularly, but their precise sizes vary from reaction tube to reaction tube. The bands do not hybridize to the specific probe, as can be seen in Figure 5B (lanes 4, 5 and 6). We have isolated and have sequenced three bands from a

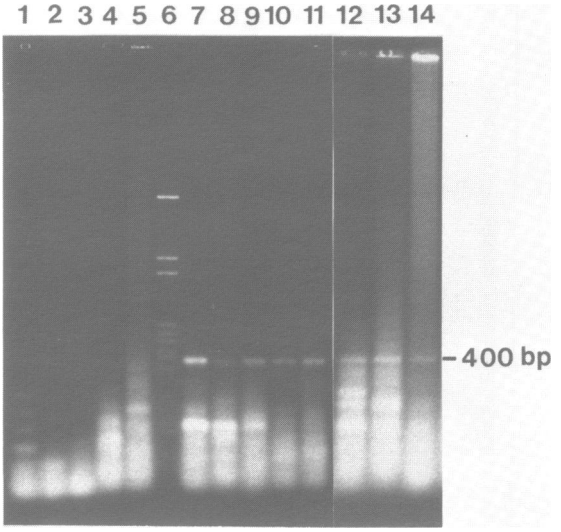

Figure 6. Detection of recombinant cells in the presence of an excess of non-recombinant (parental) cells.

Ethidium bromide stained 2% agarose gel electrophoresis of PCR amplified samples containing various numbers of non-recombinant cells in the presence or absence of recombinant HPRT-modified cells.

Lanes 1-5: Samples containing 0, 10, 10^2, 10^3 and 10^4 non-recombinant cells without any recombinant cells.

Lane 6: Size marker mix.

Lanes 7-11: Samples containing 10 recombinant HPRT-modified cells mixed with 0, 10, 10^2, 10^3, and 10^4 non-recombinant cells.

Lanes 12-14: Samples containing 5 recombinant HPRT-modified cells mixed with 0, 10^4 and 5 x 10^4 non-recombinant cells.

The 400 bp recombinant fragment, visible only when HPRT-modified cells are present in the mixture, is indicated.

typical buffer control. All three were bounded on either end by the two primer sequences. The fragments had overall lengths (including the primers) of 106 bp, 120 bp, and 169 bp. Their sequences appear to be somewhat repetitive but they bear no obvious relationships to each other, to sequences in the target locus, or to the sequences of the primers (except at their ends).

Applicability of the assay to mixtures of cells

Any assay for recombinants in a gene targetting experiment must be able to detect a few recombinant cells in the presence of non-recombinants. Previous work by Saiki et al (17) has demonstrated that small amounts of DNA containing a specific gene can be detected by PCR in the presence of a large excess of equivalent DNA that lacks the gene. We have tested, in extension of this principle, whether our assay could detect a few of the recombinant HPRT-

modified cells in the presence of an excess of the starting non-recombinant target cells. Figure 6, lanes 7-11, compares the results obtained when ten recombinant HPRT-modified cells were mixed with 0, 10, 10^2, 10^3 and 10^4 non-recombinant target cells and amplified for 50 cycles. Control reactions containing the same numbers of non-recombinant cells were also amplified without the addition of any recombinant cells (lanes 1-5). A visible band of the predicted 400 bp size is seen in all those reactions where the correctly modified recombinant cells are present, but not in any of the controls. The authenticity of the stained bands was confirmed by Southern blotting (data not shown). In Figure 6, lanes 12-14, we also show the results with five recombinant cells mixed with 0, 10^4 and 5 x 10^4 non-recombinant cells. Again the specific band is readily observed as a visible band.

The experiments illustrated in Figure 6 show that recombinant cells can be detected in the presence of up to a 10,000-fold excess of non-recombinant parental cells, without the need for hybridization to a probe. The usefulness of the method of detecting recombinants increases directly in proportion to the extent that this excess factor can be increased. We anticipate that the current level of 5 in 50,000 will be improved upon with further experience.

Absence of artefactual recombination

There is one conceivable artefact that could render the usefulness of our present assay for recombinants invalid. That would be if the presence of both parental sequences could artefactually generate recombinant molecules during the assay. Two conditions must be met for artefactual recombinant molecules to be generated. First, complementary single stranded DNA molecules corresponding to the two parental sequences must cross-hybridize to form a region of duplex DNA. Second, at least one of the two molecules involved in the duplex must have a 3′ terminus within the region of homology so that, with the other molecule acting as a template, the 3′ terminus can serve as a primer to complete synthesis of the artefactual recombinant. These conditions are most likely to be encountered when a few recombinants are being sought among a large number of cells that still carry the unmodified target locus but have the incoming DNA integrated into their genomes at random sites. We set up a worst case control experiment to see if artefacts are generated under these conditions. Genomic DNA equivalent to 10^2, 10^3, and 10^4 cells having the unmodified target locus was mixed with an equal number of DNA molecules representing copies of the incoming DNA integrated at random sites that had aborted their polymerizations at the worst place - at the 3′ end of the region of homology. The DNA used to provide these worst possible cross-primers was

the 3.3 kb incoming DNA fragment itself (middle line, Figure 4). As a positive control, we used the same mixtures in the presence of DNA equivalent to ten bona fide recombinant HPRT-modified cells. Fifty rounds of amplification were carried out. All the positive controls gave the expected specific electrophoretic band after ethidium bromide staining, and it hybridized strongly to the appropriate probe. The test samples containing only the two parental sequences completely lacked the specific band, by staining or by hybridization. Thus the assay does not generate artefactual recombinants when tested under conditions most likely to produce them.

Types of targetted modification that can be detected

There are some constraints on the types of modification we expect to be able to detect by our assay. These constraints center around the requirements that the primers be specific for the individual parental sequences, and that the recombinant fragment be of a size compatible with amplification. The choices available for meeting these requirements depend on the particular modification that is planned. Figure 7 illustrates some important parameters for two general categories of modification (24):- Ω-type (replacement), or O-type (insertion).

When the modification is of the Ω-type, there are no particular limitations on the total length of the incoming DNA. The two primers chosen to detect recombination between sequences at one end of the Ω (A and C in Figure 7A) must not be separated by more than about 2 kb, but the length of sequence homology at the other end of the Ω has no limitations. Selection of primer A is essentially unconstrained, because it is from a region of DNA not present in the incoming DNA. Primer C, on the other hand, is from a region of incoming DNA that may have an equivalent portion (C′) in the target gene, in which case C must differ sufficiently from C′ to prevent cross-priming. When C is unrelated in sequence to C′, for example when C represents the addition of unrelated DNA into the locus (as was D in Figure 4, above), this primer difference is easy to achieve. Problems may arise when C is partly related in sequence to C′. Our current experience is that discrimination in priming requires more differences than would be needed to allow an equivalent level of discrimination in Southern blots. In two different cases that we have investigated, successful primer discrimination required primers with four or more mismatches out of 20. If the targetted modification is planned to lead to a deletion, (i.e., when the chromosome is longer than the incoming DNA) it will often be acceptable to combine the deletion with the simultaneous inser-

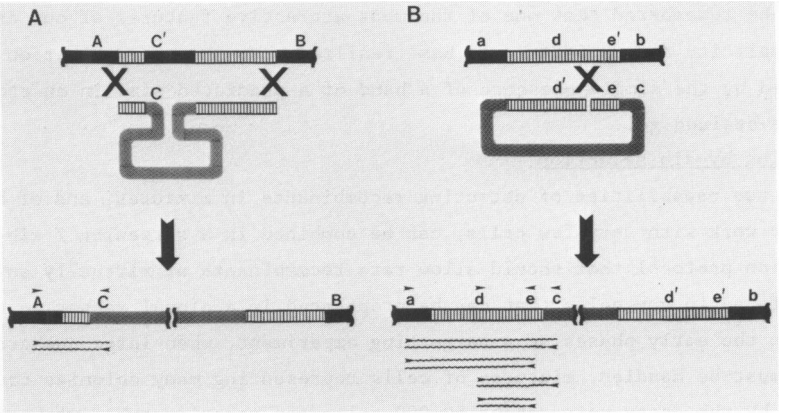

Figure 7. Parameters for the choice of primers.
Recombinants generated after targetted modification via A: an Ω -type (replacement) event; B: an O-type (insertion) event. Heavy black lines represent the target chromosome; hatched regions are sequences in the target locus and in the incoming DNA that are homologous; the stippled gray regions indicate sequences in the incoming DNA that replace or are added to the target locus. The letters are used to identify DNA regions and their corresponding primers (shown by small arrowheads) discussed in the text. The parallel wavy lines represent recombinant fragments generated by PCR amplification using the indicated primers.

tion of a small fragment of unrelated DNA, so that primer C can still be arranged to be unique to the incoming DNA.

Choice of primers to detect modifications of the O-type is straightforward when the length of DNA common to the incoming DNA and the target gene is less than the distance that can be amplified successfully. In this case the primers (a and c in Figure 7B) can be chosen from regions that have no counterparts in the other participant, but the region of homology will not then be able to exceed 2 kb. If longer stretches of homology are required, one or other or both of the primers d and e would have to be used to keep the amplified fragment within acceptable size limits. In this case, the sequences d and d' and e and e' are subject to the constraints discussed above for C and C'.

Additional discrimination can often be built into any particular system. For example a probe can be used that will hybridize to the recombinant fragment but not to sequences amplified from the target locus. Or a specific restriction site that is absent in the target can be introduced into the recombinant fragment. However, in considering these and other options, it

should be remembered that one of the most attractive features of our assay - its simplicity and rapidity - is best realized when the recombination can be detected by the simple presence of a band of a predicted size in an ethidium bromide-stained gel.

Screening by sib-selection

The two capabilities of detecting recombinants in mixtures, and of being able to work with very few cells, can be combined in a screening / sib-selection protocol that should allow rare recombinants at virtually any locus to be found in any cells that can be propagated in a clonal manner in culture. Thus in the early phases of a targetting experiment, when large numbers of cells must be handled, mixtures of cells representing many colonies can be tested by the assay using about 50,000 cells per reaction tube, with each reaction tube representing an individual primary pool consisting of 1000 - 10,000 colonies. One can thereby expect to identify primary pools in which a recombinant is present. Sub-pools, each consisting of 10-100 colonies, can then be made from some of the positive primary pools, and these can be tested by the assay to identify a sub-pool in which a recombinant is present. Individual colonies from a positive sub-pool can then be grown and tested, and a pure colony of recombinants can be identified. At none of these stages should it be necessary to expand the cell numbers to more than a hundred per colony.

CONCLUSION

In conclusion, we have developed a PCR assay for detecting the targetted modification of genes by homologous recombination that does not depend on the phenotype of the target gene or on its expression in target cells. The assay has the potential of detecting recombination between marker sequences, defined by primers, that are separated in the recombinants by up to about 2 kb. The assay can detect recombination in individual colonies of treated cells by using only a small number of these cells (10 or fewer), or it can be used to detect a similar number of recombinant cells in the presence of up to 10,000 times the number of starting non-recombinant cells.

ACKNOWLEDGEMENTS

We dedicate this paper, #2989 from the Genetics Laboratory, University of Wisconsin-Madison, to the late Joe Swoboda who constructed our automatic temperature-shift apparatus from minimal sketches. We thank Drs. Henry Erlich and Ed Sheldon, Cetus Incorporated, San Francisco, CA for their helpful advice

on PCR amplification in the early stages of our work, and Dr. Tom Doetschman for HPRT$^+$ and HPRT$^-$ embryonic stem cells. The work was supported by NIH grants GM20069 and HL37001 (to O. S.) and GM37567 to Dr. Nobuyo Maeda.

REFERENCES

1. Folger, K.R., Thomas, K., and Capecchi, M.R. (1984) Cold Spring Harbor Symp. Quant. Biol. 49, 123-138.
2. Lin, F.-L., Sperle, K., and Sternberg, N. (1984) Cold Spring Harbor Symp. Quant. Biol. 49,139-149.
3. Smithies, O., Koralewski, M.A., Song, K.-Y., and Kucherlapati, R. (1984) Cold Spring Harbor Symp. Quant. Biol. 49, 161-170.
4. Smith, A.J.H., and Berg, P. (1984) Cold Spring Harbor Symp. Quant. Biol. 49, 171-181.
5. Lin, F.-L., Sperle, K., and Sternberg, N. (1985) Proc. Natl. Acad. Sci. USA 82, 1391-1395.
6. Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A., and Kucherlapati, R. (1985) Nature 317, 230-234.
7. Rubnitz, J., and Subramani, S. (1986) Mol. Cell. Biol. 6, 1608-1614.
8. Thomas, K.R., Folger, K.R., and Capecchi, M.R. (1986) Cell 44, 419-428.
9. Song, K.-Y., Schwartz, F., Maeda, N., Smithies, O., and Kucherlapati, R. (1987) Proc. Natl. Acad. Sci. USA 84, 6820-6824.
10. Thomas, K.R., and Capecchi, M.R. (1987) Cell 51, 503-512.
11. Doetschman, T., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S., and Smithies, O. (1987) Nature 330, 576-578.
12. Saiki, R.K., Scharf, S., Faloona, F., Mulliswork, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985) Science 230, 1350-1354.
13. Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986) Cold Spring Harbor Symp. Quant. Biol. 51, 263-273.
14. Lee, M.-S., Chang, K.-S., Cabanillas, F., Freireich, E.J., Trujillo, J.M., and Stass, S.A. (1987) Science 237, 175-178.
15. Mullis, K.B, and Faloona, F.A. (1987) Methods in Enzymol. 155, 335-350.
16. Kogan, S.C., Doherty, M., and Gitschier, J. (1987) N. Eng. J. Med. 317, 985-990.
17. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, J.T., Mullisk, K.B. and Erlich, H.A. (1988) Science 239, 487-491.
18. Chien, A., Edgar, D.B., and Trela, J.M. (1976) J. Bacteriol. 127, 1550-1557.
19. Kaledin, A.S., Slyusarenko, A.G., and Gorodetskii, S.I. (1980) Biokhimiya 45, 644-651.
20. Rollo, F., Amici, A., and Salvi, R. (1988) Nucleic Acids Res. 16, 3105-3106.
21. Southern, P.J., and Berg, P. (1982) J. Mol. Appl. Genet. 1, 327-341.
22. Doetschman, T., Maeda, N., and Smithies, O. (1988) Proc. Natl. Acad. Sci. USA. In press.
23. Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1986) Nature 324, 163-166.
24. Gregg, R.G., and Smithies, O. (1986) Cold Spring Harbor Symp. Quant. Biol. 51, 1093-1099.