Nucleic Acids Research, Vol. 19, No. 19 5275–5279

Genomic fingerprinting using arbitrarily primed PCR and a matrix of pairwise combinations of primers

John Welsh and Michael McClelland*
California Institute of Biological Research, 11099 North Torrey Pines Rd, La Jolla, CA 92037, USA

Received July 9, 1991; Revised and Accepted September 9, 1991

ABSTRACT

Polymorphisms in genomic fingerprints generated by arbitrarily primed PCR (AP-PCR) can distinguish between slightly divergent strains of any organism. Single oligodeoxyribonucleotide (oligo) primers have been used to generate such fingerprints, with the same primer being present at the 5′ end of both strands for every PCR product. We used three arbitrary oligos, individually and in pairs, to generate six different genomic fingerprints of the same mouse genomic DNAs. Fewer than half of the products in genomic fingerprints generated using the oligos in pairs were the same as those produced by AP-PCR using one of the three oligos alone. Thus, a few oligos could be used in a very large number of single and pairwise combinations, each producing a distinct AP-PCR fingerprint with the potential to identify new polymorphisms. For example, 50 oligos can be used in a matrix of pairwise combinations to produce 2,500 fingerprints, in which at least half the data can be expected to be unique to each pair. We demonstrate this principle by using two oligos, alone and together, to generate three sets of fingerprints and map thirteen polymorphisms in the C57BL/6J x DBA/2J set of recombinant inbred mice.

INTRODUCTION

Simple and reproducible fingerprints of any complex genome can be generated using single arbitrarily chosen primers and the polymerase chain reaction (AP-PCR) (1,2,3,4). Sequence polymorphisms detected by genomic fingerprinting can be mapped genetically or used in phylogenetic and population studies (3,4, JW and MM, manuscript in prep.). Since AP-PCR can be semi-automated, the amount of information that could be generated is enormous.

A single oligodeoxyribonucleotide (oligo) of 10 to 34 base pairs or more is sufficient for the AP-PCR reaction and until now only single oligos have been used (1,2,3,4). Here, we show that such oligos can be used not only individually, but also in pairwise combinations. It might reasonably be expected that 50% of the PCR products in an experiment with two oligos would have the *same* primer at each 5′ end, while the remaining 50% would contain *different* primers, one at each 5′ end. Fifty oligos can be used individually or in pairwise combinations to produce 2,500 distinct genomic fingerprints. Thus, the information content contained in fingerprints produced by 50 oligos used individually and in pairwise combinations would be at least equivalent to the information content in 1,250 fingerprints using each oligo individually.

Here we show that the number of products in AP-PCR fingerprints generated by pairs of oligos does not change substantially when compared to the number produced by individual oligos, although the fingerprint pattern changes substantially. The average molecular weight of the sequences amplified by pairs of oligos is lower than when only a single oligo is used. Somewhat surprisingly, in the pairwise reaction more than half the AP-PCR products are different from those produced by either oligo alone. We demonstrate the use of oligos in pairwise combinations to detect and map sequence polymorphisms in the mouse.

METHODS

Strains

DNA from *Mus* strains AKR/J, C3H/HeJ, C57BL/6J, C57L/J, DBA/2J, MEV/1Ty, SWR/J. and *M. castaneus* strain CAST/Ei was supplied by Dr Benjamin Taylor at the Jackson Laboratory.

Primers

Oligodeoxyribonucleotides were obtained from Genosys, Houston, TX.
KA 5′ CTTGTACGCGTGTGCGAC.
KB 5′ CCTACACGCGTATACTCC.
KM 5′ CTTGCGCGCATGTACATGAC.
KR 5′ CCAAGTCGACATGGCACRTGTATACATA-YGTAAC.
KX 5′ CTTGCGCGCATACGCACAAC.

These oligos do not match any mouse sequences in the GENBANK database.

AP-PCR amplification

10 μl reactions were prepared using 0.025 units of Taq polymerase and 1X Taq polymerase buffer (Stratagene) adjusted to 4 mM $MgCl_2$, 0.2 mM of each dNTP, 10 μM of one oligo

* To whom correspondence should be addressed

or 5 μM each of two oligos and DNA at various concentrations as indicated in the figure legends. A high Mg^{2+} concentration was selected to enhance the stability of primer/template interactions. The reaction was overlaid with oil and and subjected to two cycles through the following temperature profile: 94°C for 5 min to denature, 40°C for 5 min for low stringency annealing of primer and 72°C for 5 min for extension. This temperature profile was followed by ten high stringency cycles: 94°C for 1 min, 60°C for 1 min and 72°C for 2 min. At the end of this reaction, 90 μl of a solution containing 2.5 units of Taq polymerase, 1×Taq buffer (with 1.5 mM $MgCl_2$), 0.2 mM dNTPs and 5 μCi α-[^{32}P] dCTP was added and the high stringency cycles were continued for an additional 30 rounds (2).

Autoradiography

4 μl of 80% formamide, with dye, was added to 2 μl of each PCR sample. The samples were heated to 85°C for 3 min and

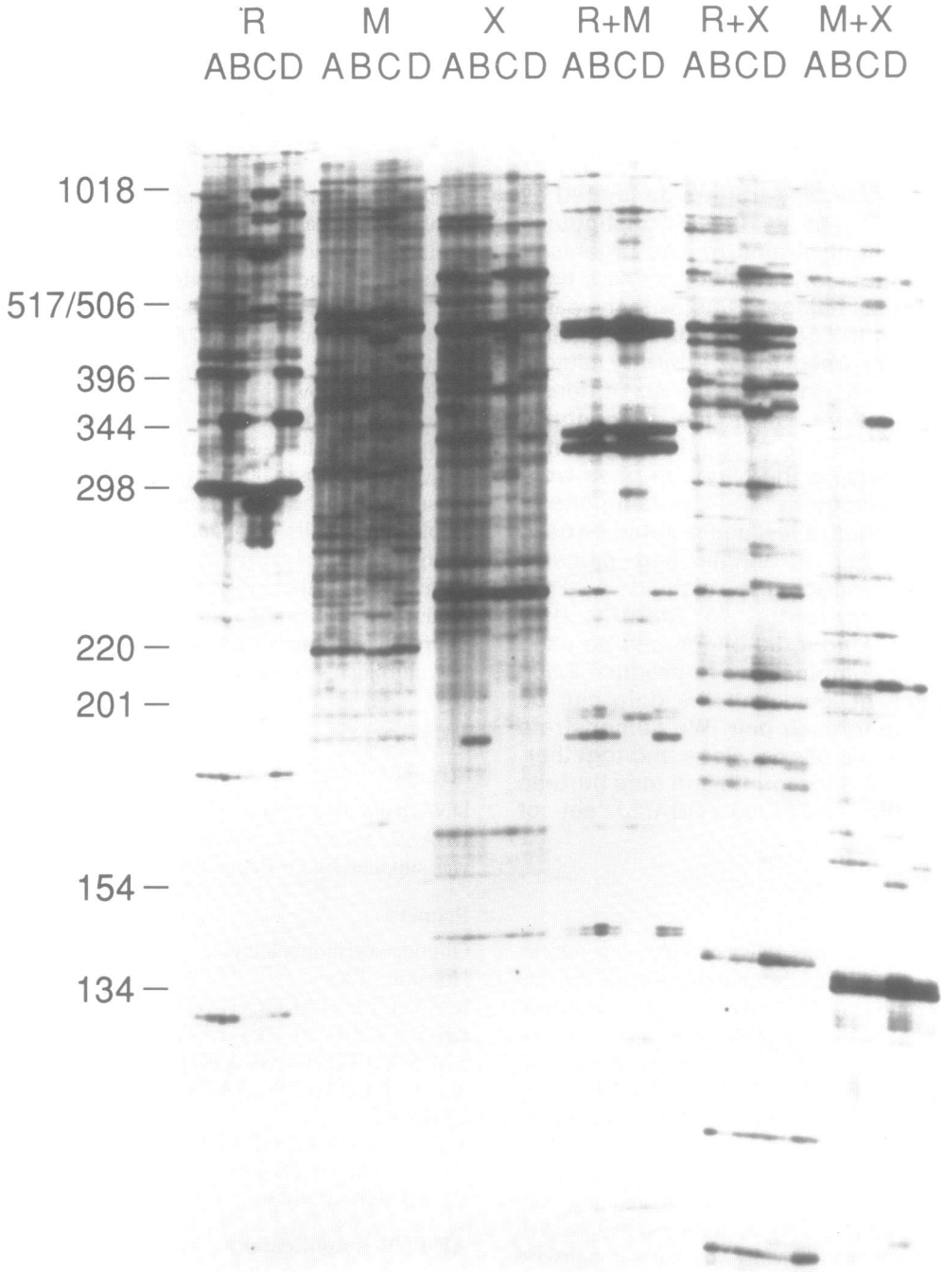

Figure 1. Pairwise oligos. The mouse strains (A) C57BL/6J, (B) DBA/2J, (C) CAST/Ei (D) MEV. were fingerprinted by AP-PCR a two concentrations of genomic DNA (10 ng on the left and 50 ng on the right) for each strain. Three oligos KM, KR or KX, were used either alone or in pairs. The fingerprints were resolved on a 5% denaturing polyacrylamide gel and autoradiographed. Two concentrations of template are used to detect any problems with reproducibility of the AP-PCR fingerprint that may occur due to pipetting errors or problems with one of the reagents. These template concentrations were chosen after experiments that showed that 5 ng to 200 ng of mouse genomic DNA gave the most reliable results (4).

2.5 μl of each was loaded on a standard 5% acrylaminde 1×TBE 50% urea sequencing gel and electrophoresed until the xylene cyanol dye had reached the bottom of the gel.

RESULTS AND DISCUSSION

Previously, we genetically mapped polymorphisms detected by AP-PCR using single oligos in a set of C57BL/6J×DBA/2J mouse recombinant inbreds (4). We wished to increase even further the simplicity and cost-effectiveness of this mapping method. To this end, we compared the fingerprints produced by AP-PCR using single arbitrarily chosen primers with the patterns generated by pairs of oligos for several important mouse strains and determined the number of polymorphisms that could be detected between strains. These strains, C57BL/6J, DBA/2J, CAST/Ei and MEV, have diverged to various extents and several have been used to construct mapping populations, such as the C57BL/6J×DBA/2J recombinant inbred set (5).

In the first experiment (**figure 1**), we used three oligos KM, KR, and KX either alone or in the three possible pairs, KM+KR, KM+KX and KR+KX, to fingerprint genomic DNAs from the C57BL/6J, DBA/2J, CAST/Ei and MEV strains. The average length of products produced by AP-PCR when two oligos were used as primers was, in all cases, shorter than that when only one oligo was used. The average length of the most prominent AP-PCR products in figure 1 was 450 base pairs for the three oligos used individually and 370 base pairs when the oligos were used in combination. This result was expected because the average distance between adequate matches in opposite strands in the genome is less when two oligos are used as primers than when only one oligo is present. Also, smaller PCR products might be expected to amplify more efficiently than larger molecules.

One might expect that half of the AP-PCR fingerprint produced by using a pair of oligos would also be found in one or the other of the patterns generated by each oligo used alone. Surprisingly, in these experiments the overlap was somewhat less. To quantitate this phenomenon, the most prominent AP-PCR products were compared for C57BL/6J in the six fingerprints. Of 49 bands examined from the experiments using a single oligo, at most 17 (34%) bands were present in the pairwise experiments. Of 52 bands examined from the pairwise experiments, 17 (32%) were present in the experiments with single oligos. One possible explanation for this observation is that the smaller products in the AP-PCR fingerprints using two primers together were competing more effectively during extension against the longer products produced by either primer alone. Thus, products

Table 1. Pairwise oligos map new polymorphisms. The C57BL/6J×DBA/2J recombinant inbred set (BXD RIs) were scored using the KA, KB and KA+KB oligos. One of the 'A+B' polymorphisms, KAB2, was the same as KB3. KA3 and KA4 are not shown because only 20 recombinant inbreds could be scored reliably. The polymorphisms are named for the oligo(s) and order on the gel. However, consistent with standard nomenclature, these will also be called *D14McC1*, *D11McC1*, *D4McC1*, *D13McC1*, *D4McC2*, *D3McC1*, *D2McC2*, *D8McC1*, *D13McC1*, *D7McC1*, *D19McC1*, *D7McC2 and D16McC1*, respectively. The previously published polymorphisms KR310, KR185/175, KR115 and KR235 (4) become *D10McC1*, *D10McC2*, *D12McC1*, *and D2McC1*, respectively. **SDP** is the strain distribution pattern. **Chr** is the chromosomal location. **Locus** lists an example of a previously mapped locus with a segregation pattern that most closely matches that of the AP-PCR polymorphism. **Lk** (linkage) gives that number of lines in which the AP-PCR polymorphism does not segregate with the nearest known loci versus the total number of BXD lines that were scored.

SDP	KA1	KA2	KAB1	KAB2	KAB3	KAB4	KB1	KB2	KB3	KB4	KB5	KB6	KB7
Chr	14	11	4	13	4	3	2	8	13	7	19	7	16
Locus	Es-10	Zfp-3	Ahd-1	As-1	Ifa	Evi-1	Ly-24	Defcr	As-1	Hbb	Xmv-18	Svp-2	Pmv-35
Lk	2/24	0/26	0/25	1/25	1/25	5/25	4/25	0/25	1/25	0/26	0/26	0/24	1/24
BXD													
1	0	1	0	0	1	1	0	0	0	0	0	1	0
2	1	0	1	0	0	1	0	0	0	0	1	0	0
5	0	1	1	1	0	0	1	1	1	1	0	1	0
6	1	0	1	1	0	0	0	0	1	0	0	0	0
8	1	1	0	1	0	1	0	0	1	0	0	1	1
9	1	1	1	0	1	0	1	0	0	1	1	1	0
11	1	1	1	0	0	0	1	1	0	0	0	1	1
12	0	1	0	0	0	1	1	0	0	1	1	1	0
13	0	0	0	0	1	1	1	1	0	1	0	1	0
14	1	1	0	0	0	1	1	1	0	1	1	0	0
15	1	0	0	0	1	1	0	0	0	1	1	1	1
16	1	0	0	0	0	0	1	1	0	0	1	1	1
18	0	1	0	1	0	1	1	1	1	1	1	1	0
19	1	0	1	1	1	1	1	1	1	0	0	1	0
20	0	1	1	0	0	0	1	0	0	0	1	1	1
21	0	0	0	0	1	1	1	1	0	0	1	1	1
22	0	1	1	1	1	0	0	1	1	0	1	1	0
23	1	1	0	0	0	0	1	0	0	0	0	1	0
24	0	1	1	0	1	0	1	1	0	1	1	1	1
25	1	1	1	1	1	0	0	0	1	1	1	1	1
27	0	1	1	0	1	1	0	0	0	0	0	1	1
28	0	0	0	0	?	?	0	0	0	1	1	0	0
29	1	0	1	1	0	1	1	1	1	0	1	?	1
30	1	1	0	0	1	1	1	1	0	0	1	1	0
31	0	1	1	0	0	1	0	1	0	1	1	0	0
32	1	0	0	0	0	0	1	1	0	1	0	0	1
B6	0	1	1	1	0	0	0	0	1	0	1	1	0
DBA	1	0	0	0	1	1	1	1	0	1	0	0	1
F1	1	1	1	1	1	1	1	1	1	?	1	1	1

containing both oligos were more prevalent in the final pattern. Alternatively, PCR products that use a single primer can produce hairpins during renaturation, possibly putting them at a disadvantage relative to PCR products made using two primers, which cannot produce hairpins. Whatever the reason, since pairwise combinations give patterns that differ by more than 50% from those produced by either primer alone, more than 50% of the polymorphisms detected between two DNAs using pairwise oligos should be different from those detected using each oligo alone. **Figure 1** also shows that AP-PCR with pairs of oligos often have a lower background than when single arbitrary selected oligos are used.

To show that AP-PCR using pairwise combinations of oligos can be used for genetic mapping, we employed two oligos, KA and KB, to produce three AP-PCR fingerprints on a population of mouse recombinant inbreds (6). KA gave four polymorphisms in the C57BL/6J×DBA/2J recombinant inbred set (BXD RIs) (**table 1**). KB gave seven polymorphisms. KA and KB together gave four polymorphisms. Three polymorphisms detected by KA+KB were not found by KA or KB oligos when used alone. One KA+KB polymorphism (KAB2) was the same as a polymorphism mapped using the KB oligo alone (KB3). All four KA+KB polymorphisms were different sizes and had a different segregation pattern compared to those detected by the KA oligo alone. Of the fifteen polymorphisms observed in this experiment three were length polymorphisms (e.g. **figure 2**, polymorphism B3) and twelve were scored by the presence or absence of an AP-PCR product (e.g **figure 2**, polymorphisms B6 and B7).

The BXD set is already well mapped by RFLP analysis. There are about 400 unique mapped SDPs in the database out of about 1600 possible SDPs likely in the 26 lines of the BXD recombinant inbreds (B. Taylor, pers. comm.). Six of 13 polymorphisms

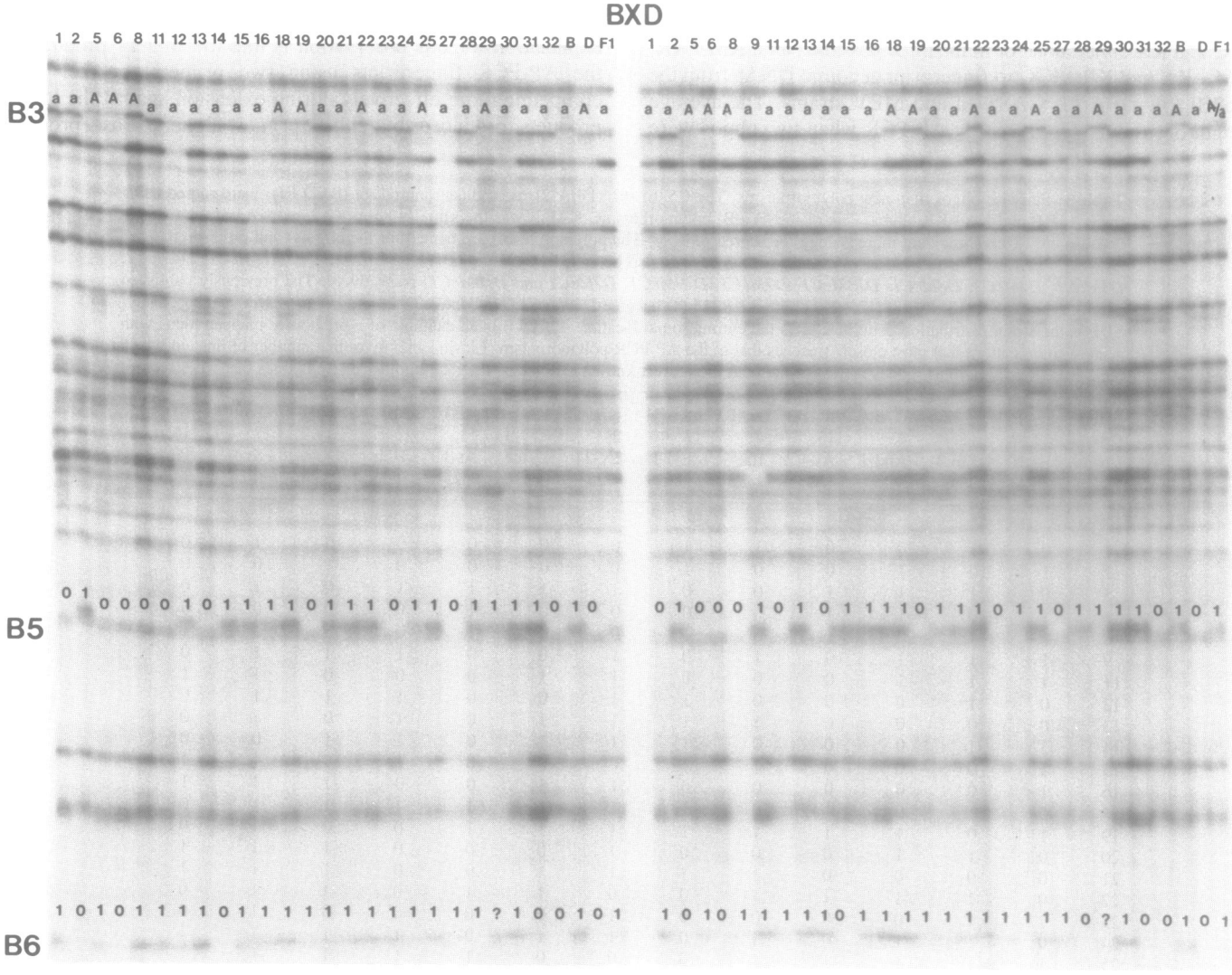

Figure 2. Example of segregating +/− and length polymorphisms using the KB oligonucleotide. Numbers at the top of the figure designate C57BL/6J×DBA/2J (BXD) recombinant inbred lines. Note the reproducibility of the two separate experiments performed with different template concentrations, 10 ng on the right and 50 ng on the left. B3 is a length polymorphism whereas B5 and B6 are scored by the presence or absence of an AP-PCR product. The figure is an autoradiograph of a small region of a 5% denaturing polyacrylamide gel. The B3 polymorphism was also present in the experiment with the KA+KB pair. 'a' and 'A' represent co-dominant length polymorphisms. The polymorphism B5 can only be scored in one set of data for BXD 8, 9, 27, and 29. B6 can only be scored in one set of data for BXD 8, 9, 24 and 29. We always perform experiments at least two template concentrations to minimize mis-scoring due to poor resolution or sporadic products that occasionally occur for a few samples.

perfectly cosegregated with a different one of the 400 previously observed SDPs. Seven polymorphisms presented here and nine of ten polymorphisms in previous (4) and unpublished AP-PCR mapping experiments appear to mark regions of the chromosome for which polymorphisms had not previously been identified. Assignments that required double cross-over events, which are unlikely, were predicted for only two of 23 AP-PCR polymorphisms we have mapped so far, including KB7, (*D16McC1*). In these two cases it is possible that we incorrectly scored one or two recombinant inbreds.

Since a set oligos can be used in a very large number of pairwise combinations, a very large number of different patterns can be produced, each of which shares less than 50% similarity with any other. For instance, 50 oligos can be used in 2,500 different combinations, either individually or in pairs. On average, each oligo should reveal four or more polymorphisms in the BXD RI set, of which two or more will be unique to that pair of oligos. This very conservative estimate indicates that 5000 polymorphisms could be mapped in the BXD RI set using only 50 oligos. Furthermore, one may be able to increase the likelihood that products unique to the pairwise fingerprint will be polymorphic by using pairwise combinations of those oligos that, when used individually, yield the most polymorphisms.

The principle of a combinatorial use of oligo pairs applies to genomic fingerprinting of any organism. Using the *same set* of 50 oligos that we could use for mapping in the mouse, a genetic mapping project could begin on the genomes of organisms as diverse as frog, fish, insect or plant, as soon as a suitable mapping population was available. Of course, the number of oligos necessary for mapping the genome of an organism is determined by the size of the genome, the level of recombination, and the desired accuracy. Depending on the application, one could choose 10 oligos, which could be used in 100 combinations, or even 100 oligos in 10,000 combinations.

Population genetics data can be produced by AP-PCR fingerprinting of a group of strains from the same species (2). A very small number of oligos in pairwise combinations could be used to generate a great deal of phylogenetically informative data.

In these and other experiments, we have used oligos that are 12 to 34 bases in length, but the same phenomenon applies equally well to other AP-PCR methods, such as RAPD, that use 10-base oligos (3, Nadeau, pers. comm.). The average size of the PCR products is also reduced in our experiments with pairs of ten-base oligos relative to single ten-base oligos (B.W.S. Sobral and McClelland, unpublished data). Further, in certain cases where fingerprints can be produced by PCR between dispersed repeats (7,8,9,10), oligos may also be used in various pairwise combinations to produce different fingerprints.

ACKNOWLEDGEMENTS

We would particularly like to thank Dr. Benjamin Taylor for supplying DNAs, invaluable advice and interpreting the polymorphisms. We also thank Drs. Joe Nadeau and Kary Mullis for helpful discussions and Drs. Manuel Perucho, Rhonda Honeycutt, Bruno Sobral and Waldemar Lernhardt for critical reading of the manuscript. Our work on AP-PCR was not supported by any government agency.

REFERENCES

1. Mullis K.B. and Faloona F.A. (1987) *Methods Enzymol.* **155**:335–350.
2. Welsh J. and McClelland M. (1990) *Nucleic Acids Res.* **18**:7213–7218.
3. Williams J.G.K., Kubelik A.R., Livak K.J., Rafalski J.A., Tingey S.V. (1990) *Nucleic Acids Res.* **18**:6531–5.
4. Welsh J., Petersen C. and McClelland M. (1991) *Nucleic Acids Res.* **19**:303–306.
5. Taylor B. Recombinant inbred strains. *In* M.F. Lyon and A.G. Searle (eds.s); *Genetic Variants and Strains of the Laboratory Mouse*. Second Edition, pp773–796, Oxford Univ. Press, Oxford, 1989.
6. Swank R.T. and Bailey D.W. (1973) Science **181**:1249–52.
7. Nelson D.L., Ledbetter S.A., Corbo L., Victoria M.F., Ramirez-Solis R., Webster T.D., Ledbetter D.H. and Caskey C.T. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**:6686–6690.
8. Ledbetter S.A., Nelson D.L., Warren S.T., and Ledbetter D.H. (1990) *Genomics* **6**:475–481.
9. Welsh J. and McClelland M. (1991) *Nucleic Acids Res.* **19**: 861–866.
10. Cox R.D., Copeland N.G., Jenkins N.A. and Lehrach H. (1991) *Genomics* **10**:375–384.