Genet. Res., Camb. (1995), **65**, *pp.* 151–155 With 3 text-figures

SHORT PAPER

Survey of genomic repeat sequence-PCRs that detect differences between inbred mouse strains

PHILIP A. WOOD* AND DOUG A. HAMM

Department of Comparative Medicine, Schools of Medicine and Dentistry, University of Alabama at Birmingham, Birmingham, Alabama, USA

(*Received 18 August 1994 and in revised form 9 January 1995*)

Summary

We have developed molecular markers that distinguish between several inbred and congenic mouse strains using polymerase chain reaction (PCR) amplification of genomic DNA repeat sequences. Mouse genomic DNA, digested with four base recognition site-restriction endonucleases, was amplified by PCR using primers for the following repeat sequences: B1 (*Alu* homolog), LINE, LLR3, IAP, human *Alu* and myoglobin. Amplification products analysed by agarose gel electrophoresis and stained with ethidium bromide produced unique DNA fragments, some of which are specific for each of 12 strains tested. This method can be used for molecular analysis of the mouse genome, including genetic monitoring.

1. Introduction

Methods that distinguish genetic differences between mouse strains using polymerase chain reaction (PCR) analysis of genomic DNA would be more useful for genetic monitoring than current methods based on allozyme biochemistry because the genome itself is being tested rather than a protein product and a larger portion of the genome can be sampled. We developed PCR-based molecular markers that detect genomic differences between inbred mouse strains based on the hypothesis that specific restriction endonuclease cleavage of mouse genomic DNA would produce templates for PCR amplification of repeat sequences which are strain specific. This method extends other procedures based on PCR amplification of repeat sequences (Herman, Nadeau & Hardies, 1992; Nadeau *et al.* 1992; Woodward, Sudweeks & Teuscher, 1992) and is an advancement over molecular methods based on identification of repeat sequence using hybridization probes on Southern blots (Jeffreys *et al.* 1987). Our method employs four base recognition site restriction endonuclease digestion of mouse genomic DNA followed by PCR amplification using primers for the mouse genomic repeat sequences LINE, LLR3, B1, and IAP-1, and human *Alu* and myoglobin repeats. Our results demonstrate that this method identifies DNA sequences which are unique for individual mouse strains and potentially useful for characterization of the mouse genome, particularly genetic monitoring of mouse strains.

* Corresponding author.

2. Materials and methods

(i) *Mice*

Mice were obtained directly from the Jackson Laboratory, Bar Harbor, ME (BALB/cJ, BALB/cBy, BALB/cByJ, SJL/J, C57BL/6J) and the National Cancer Institute, Frederick, MD (C57BL/6NCr and C3H/HeNCr). Hybrids (C57BL/6J × SJL/J) were produced in our laboratory. Liver collected at necropsy was frozen at −80 °C for genomic DNA extraction. Additional DNA samples from SWR/J, CBA/J, C3H/HeJ, C57BR/cdJ, NOD/Lt, MRL/MpJ, DBA/2J, C58/J, 129/J, A/J, C57BL/10J mice were obtained from the Jackson Laboratory Mouse DNA Resource.

(ii) *Genomic DNA preparation*

Genomic DNA was prepared by standard methods using proteinase K digestion of tissue and phenol/chloroform/isoamyl alcohol extractions. Briefly, 50–100 mg of tissue placed into 100 μl of solution I (10 mM Tris-Cl pH 7·4, 10 mM NaCl, 10 mM EDTA) was homogenized in a microcentrifuge tube using a matching pestle. Additional solution I was added to a

https://doi.org/10.1017/S0016672300033164 Published online by Cambridge University Press

Table 1. *Primer sequences*

278 (human *Alu* repeat) CCGAATTCGCCTCCCAAAGTGCTGGGATTACAG (oriented 3′ → 5′) (Nelson *et al.* 1989)
B1 TACACAGAGAAACCCTGTCTC (oriented 5′ → 3′) (Zechner *et al.* 1991)
Long interspersed repeat (LINE) CTCCTGCCTGTCCCAGAAG (oriented 3′ → 5′) (Fanning, 1983)
LLRep3 CATCTTGTGAAAACCCACACC (oriented 5′ → 3′) (Heller, Gianola & Leinwand, 1988)
Human myoglobin repeat ('Jeffreys') GGTCTAAAGCTGGAGGTGGG (oriented 5′ → 3′) (Weller *et al.* 1984)
Intracisternal A-type particle (IAP-1) TCCTGAGATGTAAGCAATA (oriented 5′ → 3′) (Ymer *et al.* 1986)

final volume of 425 μl, then 20 μl of proteinase K 10 mg/ml in sol I) and 50 μl 20% SDS were added and the mixture was incubated in a 55 °C water bath for 12 h. DNA was extracted in 500 μl of phenol:chloroform:isoamyl alcohol (1·0:0·96:0·04) and rocked at room temperature for 15 min. After centrifugation at 7000 *g* for 5 min, the phenolic (bottom) layer was removed, and this extraction step was repeated. Chloroform:isoamyl alcohol (500 μl of 0·98:0·02) was added and the tube was rocked for another 15 min. After centrifugation at 7000 *g* for 5 min, the aqueous (top) layer was transferred to a clean tube and 0·1 vol of 3 M sodium acetate pH 4·8 was added and mixed. DNA was precipitated with 2·5 vol of 100% ethanol, transferred to a sterile microcentrifuge tube, washed with 500 μl of 75% ethanol, centrifuged at 7000 *g* and vacuum dried after careful decantation of the ethanol. The resulting pellet was dissolved in 50–100 μl of TE buffer (10 mM Tris pH 7·5, 0·1 mM EDTA) and the DNA concentration was estimated by O.D. at 260 nm.

(iii) *Restriction enzyme digestion and PCR reactions*

Restriction enzymes were obtained from commercial sources and used according to the manufacturer's instructions. Genomic DNA was digested overnight at 37 °C, followed by ethanol precipitation and dilution for the PCR reactions. The PCR reaction consisted of 0·1 μg of digested genomic DNA as template, primers at 15 pmol each, 0·20 mM deoxynucleotides, and 2·5 U *Taq* polymerase (Amplitaq, Perkin/Elmer, Norwalk, CT) in buffer of 50 mM KCl, 10 mM Tris (pH 8·8), 1·5 mM MgCl, 0·1% Triton X-100. PCR conditions were: 5 min initial denaturation at 95 °C, 32 cycles of 1 min at 94°, 1 min at 55°, and 4 min at 72°, followed by 10 min at 72° and a 4° soak (until refrigerated) in a Perkin-Elmer Thermal Cycler (Norwalk, CT). Primers were a human *Alu* repeat, designated 278 (Nelson *et al.* 1989); the mouse *Alu* homolog B1 (Zechner *et al.* 1991); a long interspersed repeat (LINE) (Fanning, 1983); the LLRep3 (Heller, Gianola & Leinwand, 1988); a human myoglobin repeat ('Jeffreys') (Weller *et al.* 1984); and the intracisternal A-type particle (IAP-1) (Ymer *et al.* 1986). Primer sequences are shown in Table 1. Amplified DNA fragments were separated by electrophoresis in a 3·5% agarose gel with ethidium bromide using TAE buffer. Molecular weight (MW) standards were *Hin*d III digested lambda DNA combined with pGEM markers (Promega, Madison, WI).

3. Results and discussion

In these experiments we tested the hypothesis that we could generate unique DNA fragments useful as specific molecular markers for inbred mouse strains by using restriction enzyme predigestion of mouse genomic DNA followed by PCR using genomic repeats as primer sites. Results of *Hae* III predigestion of mouse genomic DNAs followed by PCR with a combination of the human *Alu* repeat primer 278 and the mouse LINE primer, are shown in Fig. 1. These fragments represent strain-specific sequences generated from priming sites located between restriction enzyme cleavage sites in genomic DNA. A complex pattern of PCR products is produced ranging in size from 1·2 kb to ≈ 150 bp. Even though some of these mice have common ancestral origins (Atchley & Fitch, 1991), many could be distinguished by unique DNA patterns. For example, a PCR product (denoted by the arrow at ≈ 290 bp in Fig. 1) was present in the related strains C58/J and C57BR/cdJ strains, but absent in C57BL/10J. Fragments (denoted by the bottom arrow below the 179 bp marker in Fig. 1)

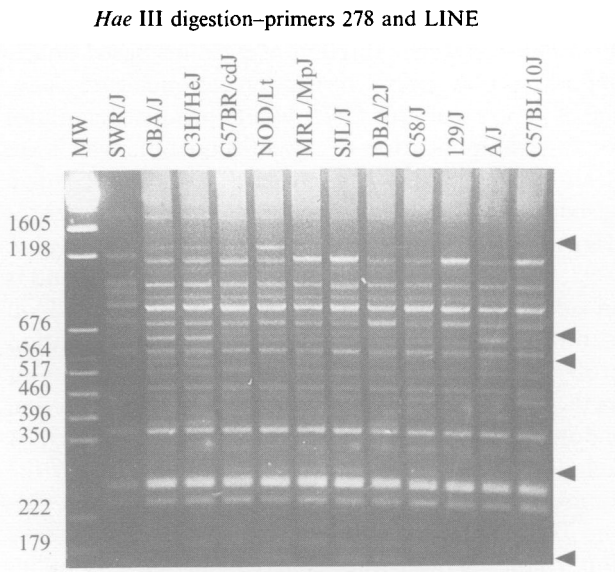

Fig. 1. PCR analysis of various inbred strains using *Hae* III digestion with primers 278 and LINE. Molecular weight markers (MWM) are *Hin*d III digested λ and pGEM markers with sizes given in the margin.

https://doi.org/10.1017/S0016672300033164 Published online by Cambridge University Press

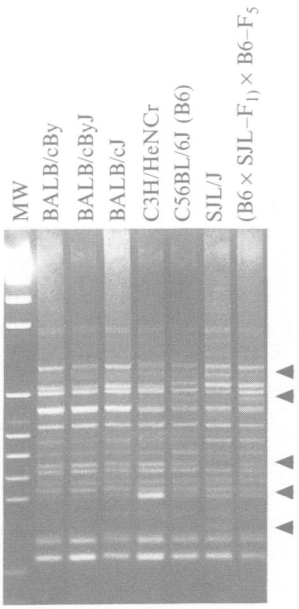

Fig. 2. PCR analysis of various inbred strains using *Taq* I digestion – primers LINE + 278. Arrows indicate PCR fragment polymorphisms among the strains. MWM are same as in Fig. 1.

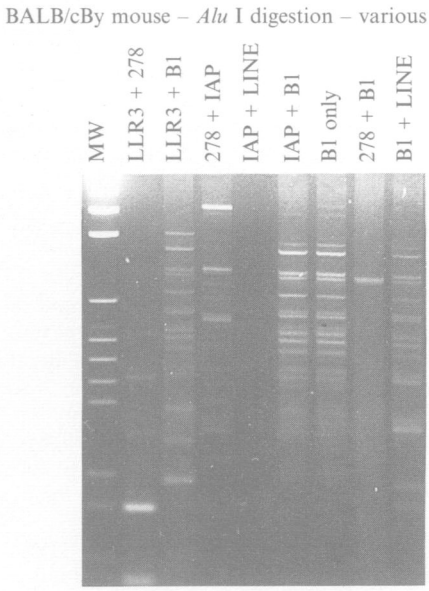

Fig. 3. PCR analysis of a BALB/cBy mouse using *Alu* I digestion with various primers and combinations. MWM are same as in Fig. 1.

which are present in NOD/Lt, SJL/J, and DBA/2J, were absent from all other strains tested. Fragments that varied among inbred strains included not only the presence or absence, or change in size, but also intensity. We found a completely different pattern by changing the 4 base restriction enzyme from *Hae* III to *Taq* I (Fig. 2) with essentially the same size fragments being produced. *Taq* I digestion of C3H/HeNCr DNA resulted in a very intense fragment seen at ≈ 350 bp (Fig. 2). We were unable to detect differences between the three BALB/c sublines analysed, despite a known deletion mutation (Hinsdale, Kelly & Wood, 1993) detectable with gene specific primers (Wood, Hinsdale & Kelly, 1993) for short-chain acyl-CoA dehydrogenase in BALB/cByJ mice. We were unable to detect PCR evidence of genetic contamination by SJL/J in a five-generation backcross of a C57BL/6J × SJL/J F_1 with C57BL/6J (Fig. 2). Both parental strain patterns were seen in the F_1 crosses (data not shown) using *Taq* I predigestion and PCR with primers LINE and 278. We believe that it would be difficult to find SJL/J sequences on the C57BL/6J background because of the high proportion of this genome present after many backcross generations.

We examined patterns generated by a third 4 base restriction enzyme, *Alu* I on BALB/cBy DNA using PCRs with several different repeat primers. A wide variety of patterns was generated (Fig. 3), presumably because of the wide range of frequencies of repeat priming sites in the genome. It has been estimated that the mouse B1 repeat, the human *Alu* repeat homolog, occurs ≈ 300000 times throughout the genome (Jelinek & Schmid, 1982). According to published B1 consensus sequence (Krayev *et al.* 1980; Jelinek & Schmid, 1982) there is a deleted T at the position comparable to the intact 4 base *Alu* I recognition sequence of the human *Alu* repeat. Thus, we found that *Alu* I digestion apparently did not cleave the 278 or B primer amplified regions. That is, after *Alu* I predigestion of mouse DNA used for PCR, we saw a set of discrete PCR products generated by the 278 and B1 primers. If the template region had been cleaved due to an *Alu* I site being present, very few PCR products should have been found. B1 should anneal with approximately the same frequency, but in the opposite direction as the primer 278. Using B1 in combination with other primers, we saw virtually the same pattern as that seen with B1 alone (Table 2 and Fig. 3). The 278 and B1 primers used on *Alu* I digested DNA produced a distinct single ≈ 800 bp fragment on a hazy background (Fig. 3) suggesting a common distance of ≈ 800 bp between B1 repeats because both primers should anneal to the B1 repeat, but in opposite directions. We also used primers designed to anneal to LINE repeats and the intracisternal A-type particle sequence (IAP-1 primer). LINE repeats belong to the family known as *Bam*HI long interspersed repeats in the mouse, which number ≈ 20000 per genome (Fanning, 1983), while IAP-1 is an endogenous retroviral-like sequence found at a frequency of ≈ 1000 copies per mouse genome (Ymer *et al.* 1986). IAP-1 alone and combined with LINE did not produce any fragments (Fig. 3), while IAP-1 combined with the 278 primer produced a simple pattern of five distinct fragments and additional faint bands (Fig. 3). The so-called 'Jeffreys' probe, used commonly in DNA fingerprinting, is derived from a human myoglobin repeat (Weller *et al.* 1984). A homologous sequence is found dispersed throughout the mouse genome and has been used in hybridization studies of

Table 2. *PCR results from various restriction enzyme digestions and various repeat priming sites*

Restriction enzyme	Primer site	Results
Hae III	278	Discrete bands, differences between B6 and BALB/c (not shown)
	278 + LINE	1 or more unique fragments per inbred strain (Fig. 1)
	LINE	No reaction product
	Jeffreys	No reaction product
	278 + Jeffreys	Few faint distinct fragments, no strain differences
	278 + B-1	Faint smear or no reaction
	B-1	Strong reaction, many fragments, few strain differences
	B-1 + LINE	Strong reaction, same as B-1 only
	B-1 + IAP-1	Strong reaction, same as B-1 only
	LINE + IAP-1	2 fragments
	IAP	No reaction product
	LLR3	No reaction product
	LLR3 + 278	9 distinct fragments, no strain differences
	LLR3 + B1	15 distinct fragments, no strain differences
	LLR3 + LINE	No reaction product
	LLR3 + IAP-1	No reaction product
Taq I	LINE + 278	Strong reaction, many unique fragments between strains (Fig. 2)
Alu I	278	Discrete bands (Fig. 3), few strain differences (not shown)
	278 + LINE	Very faint reaction products
	278 + IAP-1	10 distinct fragments, no strain differences (Fig. 3)
	278 + B-1	Single band ≈ 800 bp (Fig. 3)
	B-1	Strong reaction, no strain differences, ≈ 14 fragments found (Fig. 3)
	B-1 + LINE	Same as B-1 only (Fig. 3)
	B-1 + IAP-1	Same as B-1 only (Fig. 3)
	IAP + LINE	No reaction product (Fig. 3)
	278 + LLR3	4 distinct fragments (Fig. 3)
	278 + Jeffreys	Faint reaction products
	LLR3 + B1	16 distinct fragments (Fig. 3), few to no strain differences

mouse DNA (Jeffreys *et al.* 1987). PCRs with this primer alone produce no product, and in combination with primer 278 on *Hae* III digested DNA produced only a faint reaction product. The transcribed repeat LLRep3, which occurs at a frequency of ≈ 400 copies per genome (Heller, Gianola & Leinwand, 1988) in combination with 278 or B1 generated a set of distinct fragments (Fig. 3 and Table 2).

In our experience, clear, reproducible patterns were produced from *Alu* I, *Hae* III, or *Taq* I predigestion of mouse genomic DNA followed by PCR with either the 278 primer alone or in combination with the LINE primer. PCR using these same primers on undigested mouse genomic DNA generated a smear of DNA fragments (data not shown). Six base recognition sequence restriction enzymes also produced patterns with a high background smear of PCR products and few distinct fragments. Restriction endonuclease digestion of the PCR products had virtually no detectable effect, thus no useful restriction fragment length polymorphisms were found within the PCR products. Strain differences in electrophoretic patterns produced by several combinations of restriction enzyme predigestion and PCR primers could be catalogued and referenced against test samples for strain identification. Using this approach, the variations of restriction enzyme digestions, primers and combinations are virtually limitless. We conclude that the results reported here provide a starting point for the definition of useful methods for PCR analysis of the mouse genome which could provide a powerful method of genetic monitoring of inbred strains.

We thank Dr David Nelson for our original stock of primer 278. We thank the UAB Comprehensive Cancer Center Oligonucleotide Core Laboratory, directed by Dr Jeff Engler for synthesis of the primers used in this work. We thank Dr Henry Baker for his expert editorial assistance. This work was supported by NIH grant RR-00463.

References

Atchley, W. R. & Fitch, W. M. (1991). Gene trees and the origins of inbred strains of mice. *Science* **254**, 554–559.

Fanning, T. G. (1983). Size and structure of the highly repetitive BAM H I element in mice. *Nucleic Acids Research* **11**, 5073–5091.

Heller, D. L., Gianola, K. M. & Leinwand, L. A. (1988). A highly conserved mouse gene with a propensity to form pseudogenes in mammals. *Molecular and Cellular Biology* **8**, 2797–2803.

Herman, G. E., Nadeau, J. H. & Hardies, S. C. (1992). Dispersed repetitive elements in mouse genome analysis. *Mammalian Genome* **2**, 207–217.

Hinsdale, M. E., Kelly, C. L. & Wood, P. A. (1993). Null allele at bcd-1 locus in BALB/cByJ mice is due to a deletion in the short-chain acyl-CoA dehydrogenase gene and results in missplicing of mRNA. *Genomics* **16**, 605–611.

Jeffreys, A. J., Wilson, V., Kelly, R., Taylor, B. A. & Bulfield, G. (1987). Mouse DNA fingerprint analysis of chromosome localization and germ-line stability of hypervariable loci in recombinant inbred strains. *Nucleic Acids Research* **15**, 2823–2836.

Jelinek, W. R. & Schmid, C. W. (1982). Repetitive sequences in eukaryotic DNA and their expression. *Annual Review of Biochemistry* **51**, 813–844.

Krayev, A. S., Kramerov, D. A., Skryabin, K. G., Ryskov, A. P., Bayev, A. A. & Georgiev, G. P. (1980). The nucleotide sequence of the ubiquitous repetitive DNA sequence B1 complementary to the most abundant class

of mouse fold-back RNA. *Nucleic Acids Research* **8**, 1201–1215.

Nadeau, J. H., Bedigian, H. G., Bouchard, G., Denial, T., Kosowsky, M., Norberg, R., Pugh S., Sargeant, E., Turner, R. & Paigen, B. (1992). Multilocus markers for mouse genome analysis: PCR amplification based on single primers of arbitrary nucleotide sequence. *Mammalian Genome* **3**, 55–64.

Nelson, D. L., Ledbetter, S. A., Corbo, L., Victoria, M. F., Ramirez-Solis, R., Webster, T. D., Ledbetter, D. H. & Caskey, C. T. (1989). *Alu* polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. *Proceedings of the National Academy of Sciences*, USA **86**, 6686–6690.

Weller, P., Jeffreys, A. J., Wilson, V. & Blanchetot, A. (1984). Organization of the human myoglobin gene. *EMBO Journal* **3**, 439–446.

Wood, P. A., Hinsdale, M. E. & Kelly, C. L. (1993). Molecular detection of the *Bcd-1* null allele in BALB/cByJ mice by polymerase chain reaction: a simple assay for genetic monitoring. *Mouse Genome* **91**, 342–344.

Woodward, S. R., Sudweeks, J. & Teuscher, C. (1992). Random sequence oligonucleotide primers detect polymorphic DNA products which segregate in inbred strains of mice. *Mammalian Genome* **3**, 73–78.

Ymer, S., Tucker, W. Q. J., Campbell, H. D. & Young, I. G. (1986). Nucleotide sequence of the intracisternal A-particle genome inserted 5′ to the interleukin-3 gene of the leukaemia cell line WEHI-3B. *Nucleic Acids Research* **14**, 5901–5918.

Zechner, R., Newman, T. C., Steiner, E. & Breslow, J. L. (1991). The structure of the mouse lipoprotein lipase gene: a B1 repetitive element is inserted into the 3′ untranslated region of the mRNA. *Genomics* **11**, 62–76.