Identification of novel simple sequence length polymorphisms (SSLPs) in mouse by interspersed repetitive element (IRE)-PCR

John C. Detter¹*, Quan A. Nguyen¹ and Stephen F. Kingsmore¹,²

¹Department of Medicine and Center for Mammalian Genetics, University of Florida, 1600 SW Archer Road, PO Box 100221, Gainesville, FL 32610, USA and ²CuraGen Corporation, 555 Long Wharf Drive, New Haven, CT 06511, USA

Received April 28, 1998; Revised and Accepted July 20, 1998 MGD accession no. J46919

ABSTRACT

Interspersed repetitive element (IRE)-PCR is a useful method for identification of novel human or mouse sequence tagged sites (STTs) from contigs of genomic clones. We describe the use of IRE-PCR with mouse B1 repetitive element primers to generate novel, PCR amplifiable, simple sequence length polymorphisms (SSLPs) from yeast artificial chromosome (YAC) clones containing regions of mouse chromosomes 13 and 14. Forty-two IRE-PCR products were cloned and sequenced from eight YACs. Of these, 29 clones contained multiple simple sequence repeat units. PCR analysis with primers derived from unique sequences flanking the simple sequence repeat units in seven clones showed all to be polymorphic between various mouse strains. This novel approach to SSLP identification represents an efficient method for saturating a genomic interval with polymorphic genetic markers that may expedite the positional cloning of genes for traits and diseases.

Positional cloning is an important method of disease gene identification in human and in model organisms such as mouse. This approach is based on gene identification from contiguous genomic clones spanning a genetically defined interval that contains a locus of interest (1). A prerequisite of positional cloning is the creation of dense genetic and physical maps for the chromosomal region of interest. One of the most informative and widely-used types of genetic markers is the simple sequence length polymorphism (SSLP), also known as polymorphic microsatellite markers (2). Such markers are made up of serial di-, tri-, tetra- or penta-nucleotide repeats, where the number of repeat units differs among individuals/strains within a species. Microsatellite markers are easily amplified by PCR and exhibit strain-specific polymorphisms upon observation on agarose or polyacrylamide gels. However, for most regions of the mouse genome, current SSLP density is insufficient for positional cloning approaches. Conventional approaches to the identification of novel SSLPs from a genomic region, such as subcloning genomic clones and hybridization with repeat probes, are cumbersome.

Interspersed repetitive element (IRE)-PCR is a very useful method for identification of novel human or mouse sequence tagged sites (STTs) from contigs of genomic clones or somatic cell hybrids (3–6). This technique achieves selective amplification of mammalian DNA from an admixture with yeast or prokaryotic DNA, without prior knowledge of the mammalian DNA sequence, by PCR using outwardly oriented primers that are specific for mammalian high copy number repetitive elements. For example, IRE-PCR of mouse genomic DNA may be accomplished using primers designed from the 5′ and 3′ ends of the B1 repetitive element DNA sequence, which belongs to the short interspersed element (SINE) family (Fig. 1). The 130–150 bp B1 repetitive element, which is believed to be homologous to the human ‘Alu’ repeat, is present at a copy number of 130 000–180 000 per haploid mouse genome (7), and frequently is located at intron–exon junctions within genes (8).

We now report that B1–B1 IRE-PCR can also be exploited to identify novel polymorphic SSLPs from genomic clones containing mouse DNA. Genomic DNA from four different mouse chromosome 14 YAC clones was isolated as part of a positional cloning effort and used as the template for three separate IRE-PCR reactions. The DNA was amplified with a 5′ B1 reverse primer alone, a 3′ B1 forward primer alone and a combination of 5′ B1 reverse and 3′ B1 forward primers together (Fig. 1). The PCR master mixes contained (for one 50 µl reaction): 50 ng of yeast genomic DNA (containing a YAC), 1 µM of each primer, 200 µM of each dNTP, 1× PCR buffer and 2.5 U of Taq DNA polymerase (Boehringer Mannheim, Indianapolis, IN). The PCR amplifications were performed in a PTC-100 Programmable Thermal Controller (MJ Research, Watertown, MA). The conditions involved an initial 2 min denaturation at 95°C; then 33 cycles of 95°C for 20 s, 55°C for 20 s and 72°C for 90 s; and finally a 10 min polishing extension at 72°C. Upon amplification, an aliquot of each reaction was electrophoresed on a 2% agarose gel to ensure amplicon (PCR product) complexity of at least 10 readily identified bands. Using these conditions, amplicons varied in size between 200 and 2500 bp. Following purification of the remaining 45 µl of PCR reaction

*To whom correspondence should be addressed. Tel: +1 352 392 6865; Fax: +1 352 392 8483; Email: jдетter@nervm.nerdc.ufl.edu
from each Y AC were admixed and 6 (Wizard kit, Promega, Madison, WI), the three IRE-PCR products from which they had been derived. Testing DNA from several single amplicon of expected size upon PCR using the Y AC clone clones to permit specific amplification of these loci. All gave a designed from unique sequences flanking the repeat units in six chromosomes 14 Y ACs; data not shown). Specific primers were simple sequence repeats were identified in each of the four contained simple sequence repeats (at least two clones containing (Perkin-Elmer, Foster City, CA). 310 Genetic Analyzer and appropriate cycle sequencing reagents were end-sequenced with M13 vector primers using an ABI PRISM DNA was isolated (Wizard Plus Miniprep, Promega) and inserts of these polymorphisms to map to the same region of mouse chromosome 14 as the Y ACs (data not shown). Prior to these experiments the four chromosome 14 Y ACs were known to chromosome 13, were used as template for B1–B1 IRE PCR. Of the 20 resultant amplicons that were cloned and across the mouse genome, four Y ACs from another mouse autosome, chromosome 13, were used as template for B1–B1 IRE PCR. The first column lists the novel polymorphic markers identified with the described method. The second column lists the forward and reverse primers (5′→3′) to amplify each polymorphic marker. The third column lists the repeat units in six clones to permit specific amplification of these loci. All gave a single amplicon of expected size upon PCR using the YAC clone from which they had been derived. Testing DNA from several mouse strains revealed all six STSs to be polymorphic (Table 1). The apparent association of SSLPs with B1 repetitive elements was not anticipated. However, the finding that sequences rich in di-, tri-, tetra- or penta-nucleotide repeats found in each polymorphic marker. The last column lists the product sizes (in bp) of each polymorphic marker when using B6 (C57BL/6J), CAST (CAST/Ei), PWK (wild-derived Mus musculus musculus) and DBA (DBA/2J) mouse strain DNA as template. B6 product sizes were based upon DNA sequence analysis.

ACKNOWLEDGEMENTS

This work was supported by grants from the American Cancer Society, the Arthritis Foundation, GlaxoWellcome Research and Development Ltd, the Howard Hughes Medical Institute and the National Institutes of Health (AI39651).

REFERENCES


The table below lists the novel simple length polymorphisms identified with the described method. The forward and reverse primers (5′→3′) were used to amplify each polymorphic marker. The product sizes (in bp) of each polymorphic marker when using B6 (C57BL/6J), CAST (CAST/Ei), PWK (wild-derived Mus musculus musculus) and DBA (DBA/2J) mouse strain DNA as template. B6 product sizes were based upon DNA sequence analysis.

<table>
<thead>
<tr>
<th>Novel SSLP</th>
<th>Primer Sequence (5′ to 3′)</th>
<th>Repeat Type</th>
<th>B6 Product Size (bp)</th>
<th>CAST</th>
<th>PWK</th>
<th>DBA</th>
</tr>
</thead>
<tbody>
<tr>
<td>D14S1k1</td>
<td>Forward + GAGAAGAACCCTCGTCCTGCAAC</td>
<td>CAAA, AGGG, &amp; AG 239 225* 240* 240* 225* ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D14S1k2</td>
<td>Reverse - CTCAAGAGATGCTGCAAGTACG</td>
<td>CAAA 224 250* 225* ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D14S1k3</td>
<td>Reverse - GGACGCAGGCTATGTAATGC</td>
<td>CCAA 141 160* 140* 150*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D14S1k4</td>
<td>Forward + GTGCTCGAGCGCAAGCAACG</td>
<td>GT &amp; GA 255 255* 270* ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D14S1k5</td>
<td>Reverse - GGA4CTCTCAGAAAGGATGAC</td>
<td>CA 109 135* 110* 125*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D14S1k6</td>
<td>Reverse - CTTGGGATGTCTTTCGCTGC</td>
<td>ATGGTATAGTAACTGATGGCC 148 170* 160* 140*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*CAST, PWK and DBA product sizes are estimated, based on agarose gel electrophoresis analysis. ND, not determined. MGD accession number J46919.

(Wizard kit, Promega, Madison, WI), the three IRE-PCR products from each YAC were admixed and 6 µl of each pool was cloned en masse (TA cloning kit, Invitrogen, San Diego, CA). Twenty-two recombinant colonies were picked, grown in liquid culture, plasmid DNA was isolated (Wizard Plus Miniprep, Promega) and inserts were end-sequenced with M13 vector primers using an ABI PRISM 310 Genetic Analyzer and appropriate cycle sequencing reagents (Perkin-Elmer, Foster City, CA).

End sequence was obtained from 19 clones, 15 of which contained simple sequence repeat sequences (at least two clones containing simple sequence repeats were identified in each of the four chromosomes 14 YACs; data not shown). Specific primers were designed from unique sequences flanking the repeat units in six clones to permit specific amplification of these loci. All gave a single amplicon of expected size upon PCR using the YAC clone from which they had been derived. Testing DNA from several mouse strains revealed all six STSs to be polymorphic (Table 1). Genetic mapping in an intersubspecific backcross revealed each of these polymorphisms to map to the same region of mouse chromosome 14 as the YACs (data not shown). Prior to these experiments the four chromosomes 14 YACs were known to contain only three SSLPs. The experiments described here have identified six novel SSLPs within three of the four YACs. Table 1 lists the specific primer sequence for each marker along with the approximate amplicon size for each mouse strain tested.

In order to determine the general applicability of this method across the mouse genome, four YACs from another mouse autosome, chromosome 13, were used as template for B1–B1 IRE PCR. Of the 20 resultant amplicons that were cloned and end-sequenced, 14 contained simple sequence repeats, one of which was tested and found to be polymorphic (Table 1). The latter did map back to the region of mouse chromosome 13 from which the YAC clones were derived (data not shown). These data indicate that B1–B1 IRE-PCR is an efficient method for identification of novel mouse SSLPs. It should be noted that the remaining IRE-PCR products that may not contain simple sequence length polymorphisms are still useful as novel STSs.

The first column lists the novel polymorphic markers identified with the described method. The second column lists the forward and reverse primers (5′→3′) to amplify each polymorphic marker. The third column lists the repeat units in six clones to permit specific amplification of these loci. All gave a single amplicon of expected size upon PCR using the YAC clone from which they had been derived. Testing DNA from several mouse strains revealed all six STSs to be polymorphic (Table 1). The apparent association of SSLPs with B1 repetitive elements was not anticipated. However, the finding that sequences rich in di-, tri-, tetra- or penta-nucleotide repeats found in each polymorphic marker. The last column lists the product sizes (in bp) of each polymorphic marker when using B6 (C57BL/6J), CAST (CAST/Ei), PWK (wild-derived Mus musculus musculus) and DBA (DBA/2J) mouse strain DNA as template. B6 product sizes were based upon DNA sequence analysis.

*CAST, PWK and DBA product sizes are estimated, based on agarose gel electrophoresis analysis. ND, not determined. MGD accession number J46919.

(Wizard kit, Promega, Madison, WI), the three IRE-PCR products from each YAC were admixed and 6 µl of each pool was cloned en masse (TA cloning kit, Invitrogen, San Diego, CA). Twenty-two recombinant colonies were picked, grown in liquid culture, plasmid DNA was isolated (Wizard Plus Miniprep, Promega) and inserts were end-sequenced with M13 vector primers using an ABI PRISM 310 Genetic Analyzer and appropriate cycle sequencing reagents (Perkin-Elmer, Foster City, CA).

End sequence was obtained from 19 clones, 15 of which contained simple sequence repeat sequences (at least two clones containing simple sequence repeats were identified in each of the four chromosomes 14 YACs; data not shown). Specific primers were designed from unique sequences flanking the repeat units in six clones to permit specific amplification of these loci. All gave a single amplicon of expected size upon PCR using the YAC clone from which they had been derived. Testing DNA from several mouse strains revealed all six STSs to be polymorphic (Table 1). Genetic mapping in an intersubspecific backcross revealed each of these polymorphisms to map to the same region of mouse chromosome 14 as the YACs (data not shown). Prior to these experiments the four chromosomes 14 YACs were known to contain only three SSLPs. The experiments described here have identified six novel SSLPs within three of the four YACs. Table 1 lists the specific primer sequence for each marker along with the approximate amplicon size for each mouse strain tested.

In order to determine the general applicability of this method across the mouse genome, four YACs from another mouse autosome, chromosome 13, were used as template for B1–B1 IRE PCR. Of the 20 resultant amplicons that were cloned and end-sequenced, 14 contained simple sequence repeats, one of which was tested and found to be polymorphic (Table 1). The latter did map back to the region of mouse chromosome 13 from which the YAC clones were derived (data not shown). These data indicate that B1–B1 IRE-PCR is an efficient method for identification of novel mouse SSLPs. It should be noted that the remaining IRE-PCR products that may not contain simple sequence length polymorphisms are still useful as novel STSs.