Detection and isolation of nuclear haplotypes by PCR-SSCP

G. ORTÍ, M. P. HARE* and J. C. AVISE
Department of Genetics, University of Georgia, Athens, GA 30602, USA

Abstract
SSCP (single-strand conformational polymorphism) is used widely in the field of human biomedicine, but its potential as a population genetics tool for the recovery of nuclear gene genealogies remains to be realized. We describe and illustrate a use for SSCP in the physical isolation of nuclear haplotypes that circumvents several difficulties associated with more conventional cloning procedures. The DNA sequence can be determined directly from the isolated haplotypes and used for phylogenetic inference. SSCP provides a convenient first step toward generating nuclear genealogies for population studies.

Keywords: SSCP, allelic variation, nuclear haplotypes, gene genealogies, Limulus, Crassostrea

Received 10 September 1996; revision accepted 14 December 1996

Introduction
The concept of gene trees – the phylogeny of alleles or haplotypes for any specified stretch of DNA – has prompted phylogenetic thought and historical perspectives in population biology (e.g. Wilson et al. 1985; Avise et al. 1987; Avise 1989). However, due to both biological and technical complications, few empirical studies have assessed nuclear gene genealogies in natural populations (e.g. Hey & Klinman 1993; Bernardi et al. 1993; Palumbi & Baker 1994; Vogler & DeSalle 1994). The primary biological snag is historical intragenic recombination among homologous haplotypes, a phenomenon that requires statistical methods for detection (Sawyer 1989; Hein 1990; Maynard Smith 1992; Templeton et al. 1992) and if common can obscure the otherwise lineal histories of allelic descent. The primary technical difficulty, and the one addressed in the current paper, is the physical isolation of individual haplotypes from diploid tissue sources.

Here we describe and exemplify a general approach for isolating single-locus haplotypes from diploid organisms. The approach involves ‘single-strand conformational polymorphism’ or SSCP (Orita et al. 1989a,b), a popular technique that allows rapid detection and isolation of nucleotide polymorphisms in nondenaturing gels. Since its inception, over 1100 publications have reported the use of SSCP, mainly in biomedical research and as a clinical diagnostic tool for human disease mutations (BIOSYS, Philadelphia, PA). Only a small fraction of these studies (33 publications, < 3%) were indexed as ‘population genetics’, but most of the latter (30) were surveys of human populations, with none involving a genealogical perspective. The potential of SSCP for evolutionary biology is not yet fully appreciated (Lessa & Appelbaum 1993; Potts 1996).

Materials and Methods
DNA samples from American oysters Crassostrea virginica were procured for previous studies, and PCR primers for anonymous single-copy nuclear (scn) DNA loci were developed (Karl & Avise 1992; Hare et al. 1996). For the purposes of an example, the isolation of haplotypes from one such locus (CV-32) is presented here. To assess the efficiency of SSCP for detection and isolation of haplotypes, only individuals known to be heterozygous for a restriction site at this locus (Hare et al. 1996) were tested. Amplification by PCR was performed with 0.25 µM of each primer, CV-32.3LS (5’- CTGTACTTGTTTTCCCGTT-3’) and CV-32.7R (5’-CATTCATTTTCCTCCAAGTTTTGTT-3’), 200 µM of each dNTP, 1 mM MgCl₂, 1 x reaction buffer, and 1 unit Taq polymerase (Promega), in a final volume of 50 µL. Cycling conditions were 94 °C for 1 min, 51 °C for 45 s, 72 °C for 1 min, for 25 cycles. For verification prior to SSCP analysis, PCR products of the expected size (447 bp) were visualized in ethidium bromide-stained 2% agarose gels.

© 1997 Blackwell Science Ltd
From the horseshoe crab Limulus polyphemus (originally collected by Saunders et al. 1986), scnDNA loci were obtained by a procedure similar to that described by Sültmann et al. (1995). Briefly, randomly amplified fragments generated by the PCR with a non-specific primer (5'-CCTTATTTAAACCAACCAGCGGCA-3') were size-selected and cloned using the pGem-T vector system (Promega). Cloned inserts were sequenced from various individuals, aligned and searched for polymorphisms. Specific PCR primers were designed based on the alignment. Several scnDNA loci were obtained by the procedure (manuscript in preparation), but for the alignment. Several scnDNA loci were obtained by a procedure similar to that described by Sültmann et al. (1995). Briefly, randomly amplified fragments generated by the PCR with a non-specific primer (5'-CCTTATTTAAACCAACCAGCGGCA-3') were size-selected and cloned using the pGem-T vector system (Promega). Cloned inserts were sequenced from various individuals, aligned and searched for polymorphisms. Specific PCR primers were designed based on the alignment. Several scnDNA loci were obtained by the procedure (manuscript in preparation), but for purposes of example only the results from locus LP-1 are shown here. Specific primers for this locus (LP1s: 5'-CTTTAAGCTGCTACCCTAAC-3'; LP1a: 5'-CAGAGACGACAGTTG-3') amplify a fragment about 775 bp long. Amplification by PCR was performed with 0.4 μM of each primer, 1 mM of each dNTP, 2 mM MgCl2, 1× PCR buffer, and 1 unit of Taq polymerase, in a final volume of 50 μL. Cycling conditions were 94 °C for 1 min, 51 °C for 1 min, 72 °C for 1 min, for 29 cycles. The presence of a single clear band was verified in 2% agarose gels prior to SSCP analysis.

For both the oysters and horseshoe crabs, nonisotopic SSCP analysis was conducted on the PCR-amplified fragments following the protocol of Hongyo et al. (1993), with some modifications. Briefly, 10–20 μL of unpurified PCR product (roughly 0.5–1.6 μg of DNA) were mixed with 5 μL of denaturing/loading buffer containing 0.4 μL of 1 mM methylmercury hydroxide (Matthey Electronics, Inc., War Hill, MA), 1 μL of 15% Ficoll loading buffer (with 0.25% bromphenol blue and 0.25% xylene cyanol), and 3.6 μL of 1× TBE buffer (90 mM Tris, 92 mM boric acid, 2.5 mM EDTA). This mixture was denatured for 4 min at 85 °C and immediately chilled on ice before loading to the SSCP gel. Six to 16% polyacrylamide (39:1 acrylamide to bis-acrylamide) TBE gels (16 cm×14 cm×1.5 mm) were run with 1× TBE buffer on a vertical electrophoresis system (Fisher Biotech model VE16-1). Refrigerated water from a thermostatically controlled circulator (Brinkmann RC 20B, Lauda, Germany) was passed through a vertical cooling chamber to maintain the upper buffer and gel at a constant temperature. In a modification of the gel rig, air bubbles were introduced into the upper buffer to eliminate thermal gradients between the cooling chamber and gel. Gels were run at constant power (8–12 W) for 8–20 h, and with the upper buffer temperature set to 3–15 °C. Temperature, power, acrylamide concentration and fragment size affected the running time. Acrylamide concentration and temperature were optimized for each fragment to obtain maximum separation of allelic variants.

Gels were stained for at least 20 min with a 2 μg/mL ethidium bromide solution, and destained for 5 min with distilled water. Bands were visualized and photographed under UV light, and a small fraction of each band was excised from the gel with the tip of a 200-μL glass micropipette. These acrylamide plugs were placed individually in tubes with 50 μL of distilled water and stored at −20 °C. Prior to reamplification, the gel samples were heated to 80 °C for 10 min.

Genetic differences among the haplotypes isolated by SSCP were verified by direct sequencing, either from single stranded template generated by asymmetric PCR (Gyllensten & Erlich 1988), or from reamplified double-stranded products (fmol cycle-sequencing, Promega), and the results will be presented elsewhere in the context of phylogenetic analyses.

**Results**

The ability of SSCP to detect and isolate nuclear haplotypes is summarized in Table 1. Oyster samples (*n* = 47) from Massachusetts to Louisiana revealed 21 distinct alleles at the CV-32 locus. Figure 1 shows the effect of

<table>
<thead>
<tr>
<th></th>
<th>CV-32</th>
<th>LP-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. individuals assayed by SSCP</td>
<td>47</td>
<td>44</td>
</tr>
<tr>
<td>No. individuals showing band separation in SSCP gels*</td>
<td>37</td>
<td>26</td>
</tr>
<tr>
<td>No. unique alleles isolated</td>
<td>21</td>
<td>11</td>
</tr>
<tr>
<td>Minimum difference between alleles showing band separation</td>
<td>2 bp subst. + 1 bp-indel (lane 7, Fig. 2a)</td>
<td>1 bp subst.</td>
</tr>
<tr>
<td>Maximum difference between alleles showing band separation</td>
<td>8 bp subst. + 1 bp-indel (not shown)</td>
<td>6 bp subst. + 39 bp-indel (lane 6, Fig. 2a)</td>
</tr>
</tbody>
</table>

*Three- or four-banded pattern, allowing excision of individual haplotypes from the gel.

© 1997 Blackwell Science Ltd, Molecular Ecology, 6, 575–580
ISOLATING ALLELES WITH PCR-SSCP 577

Fig. 1 Separation of allelic variants of the CV-32 locus (447 bp) by SSCP analysis at 3 °C. Polyacrylamide gels (6–12%, bottom to top) were run at 3 °C (upper buffer temperature) for 6–15 h (bottom to top) at constant power (9 W). Each lane of the gel was loaded with PCR products from single individuals, whose genotypes were as follows (upper/lower case indicates absence/presence of a restriction site). Lane: (1) 1-kb ladder; (2) n1n1; (3) n1N1; (4) n2N2; (5) n3N1; (6) n4N3; (7) n5N1; (8) n6N4; (9) n7N5; (10) n8N6. Variation among alleles (determined by sequencing) included 1–9 bp substitutions and 1–3 indels with length variation from 1 to 4 bp.
applications empirical assessments are necessary. Our data support the idea that temperature and gel concentration are critical parameters affecting mobility shifts that allow separation by SSCP.

Discussion

Several experimental approaches and specialized genetic systems have been suggested for isolating individual haplotypes from nuclear loci (review by Avise 1994: 134). These include: (i) extraction of identical-by-descent chromosomes via controlled crosses (e.g. in Drosophila, Aquadro et al. 1986); (ii) use of haploid tissues, or of haploid phases of a haplo-diploid life cycle (e.g. Guttman & Dykhuizen 1994); (iii) use of sex-linked genetic markers (e.g. Bishop et al. 1985); and (iv) cloning of PCR products (Scharf et al. 1986). This latter approach clearly has the broadest taxonomic applicability. However, because cloning occurs through a single molecule, the possibility of nucleotide misincorporation by Taq polymerase must be considered (e.g. Keohavong & Thilly 1989; Palumbi & Baker 1994). In practice, possible Taq misincorporations either have been ignored for phylogenetic analyses (Palumbi & Baker 1994; Vogler & DeSalle 1994), or several separate clones from each individual were sequenced to distinguish allelic variation present in the population from PCR-cloning artefacts (Bernardi et al. 1993). PCR-SSCP offers a simpler and less laborious alternative to the direct cloning of PCR products because it bypasses the problem of Taq error and the need for multiple sequencing of each haplotype.

Another advantage conferred by SSCP separation (compared with the cloning of PCR products) is the ability...
to detect haplotype variants with low amplification efficiencies. When PCR amplification from a heterozygote yields a low copy number of one allele, many clones must be sequenced to isolate both haplotypes from an individual. In contrast, a mutant haplotype in SSCP gels can be detected when it comprises as little as 3% of a mixture of PCR products (Hongyo et al. 1993; Law et al. 1996). An alternative broadly applicable method for allele-specific sequencing using restriction enzymes and biotinylation has been described recently (Zhang & Hewitt 1996). Shortcomings of this method are that interallelic sequence differences must be known a priori (a special difficulty if length variation exists among alleles), and that ‘restriction cuttable’ sites for the polymorphic DNA sequences in question must be identified.

An essential characteristic sought for target loci in population genealogical analysis is the presence of substantial polymorphism in a relatively short stretch of DNA. In general, shorter fragments are easier to separate in SSCP gels and also should be less likely to have undergone intragenic recombination in recent history. Conversely, larger fragments should tend to accumulate more polymorphic sites, but SSCP separation becomes less efficient and recombination perhaps more likely. In conclusion, several advantages exist for SSCP in the detection and isolation of nuclear haplotypes for gene-genealogical studies at the population level. With physical methods for haplotype isolation now available, it will next be important to determine the extent to which rapidly evolving nuclear genes can be identified that also are relatively free of homoplasy and intragenic recombination.

Acknowledgements

D. Guttman has brought to our attention the potential of SSCP for this particular application. Research was supported by a grant from the NSF to J.C.A.

References


Gyllensten UB, Erlich HA (1989) Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQα locus. Proceedings of the National Academy of Science USA, 85, 7652–7655.


© 1997 Blackwell Science Ltd, Molecular Ecology, 6, 575–580

ISOLATING ALLELES WITH PCR-SSCP 579


Guillermo Ortí is a postdoctoral associate and Matt Hare a recent PhD graduate from the laboratory of John Avise at the University of Georgia. This work is part of a broader research program concerned with the principles and processes governing the geographic distributions of genealogical lineages (phylogeography). M. Hare is now a postdoctoral associate with S. Palumbi at Harvard University.