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

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.

Correspondence: Guillermo Ortí. Fax: +1-706-542-3910. E-mail: gorti@bscr.uga.edu

*Present address: Department of Organismic and Evolutionary Biology, Harvard University Biolabs, 16 Divinity Ave., Cambridge, MA 02138, USA.

© 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 Šultmann et al. (1995). Briefly, randomly amplified fragments generated by the PCR with a non-specific primer (5'-CCTTATATAAAGCAAACCGAGGAGCA-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 a procedure similar to that described by Saunders et al. (1995). Briefly, randomly amplified fragments generated by the PCR with a non-specific primer (5'-CTTAAACGCTTATACCCCTACG-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 x 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 a few 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 M 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 x 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 x 14 cm x 1.5 mm) were run with 1 x 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. (lane 7, Fig. 2a)</td>
<td>1 bp subst. (not shown)</td>
</tr>
<tr>
<td>Maximum difference between alleles showing band separation</td>
<td>8 bp subst. + 1 bp-indel (lane 6, Fig. 2a)</td>
<td>6 bp subst. + 39 bp-indel (lane 7, Fig. 2a)</td>
</tr>
</tbody>
</table>

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

acrylamide concentration on band separation. Homozygous individuals (e.g. lane 2), or heterozygotes whose haplotypes fail to separate (e.g. lane 9) typically show a two-banded pattern, corresponding to the pyrimidine-rich and purine-rich DNA strands. Successful isolation of haplotypes from heterozygotes is evidenced when the upper and/or lower strands separate in the gel (resulting in three- or four-banded patterns, e.g. lanes 5 and 6, respectively). Note that the lower bands of heterozygote n5N1 in lane 8 separated only at gel concentrations ≥10%. By sequencing, these alleles proved to differ by six nucleotide substitutions. Conversely, good separation of the upper bands in heterozygote n6N5 in lane 10 (four nucleotide substitutions) was obtained in 6–10% acrylamide gels but not in 12% gels. In general, optimal acrylamide gel concentrations for the CV-32 locus was determined empirically (Fig. 1) to be about 8–10%, in agreement with previous reports (Glavac & Dean 1993).

For horseshoe crabs, the primer-pair LP1s-LP1a successfully amplified a single 775-bp fragment in most individuals tested. Samples from New Hampshire to the Gulf of Mexico in Florida (n = 44) revealed 11 distinct alleles at the LP-1 locus (Table 1). Optimal conditions for the separation of alleles were obtained at 7–8 °C with 9–10% polyacrylamide gels (Fig. 2a). In lanes 4 and 6, the four-banded patterns correspond to genotypes L1S1 and L6S1 that carry a 39-bp deletion (Fig. 2b) in one of the alleles. In lanes 2 and 3, genotypes L1L4 and L1L5 (the alleles proved to differ by three nucleotide substitutions) separated at their lower strands at 8 °C, but not at 3 °C. In some cases (e.g. the heterozygote in lane 7), the lower strands separated at one temperature (3 °C) and the upper strands separated at a different temperature (8 °C). This last observation demonstrates that no single optimum set of conditions for allelic separation need exist among multiple heterozygous genotypes. Three randomly picked samples (out of 18) that did not show band separation at the LP-1 locus were confirmed to be homozygotes upon sequencing, suggesting a relatively high efficiency of this technique (Table 1).

In summary, numerous variant haplotypes at scnDNA loci from oysters and horseshoe crabs were isolated successfully by the PCR-SSCP approach. Most prior applications of SSCP report high efficiency for the detection of mutations in DNA fragments ≤200 bp-long, but fragments as large as 775 bp were isolated and sequenced in this study. Other authors (Hayashi 1992; Hayashi & Yandell 1993; Glavac & Dean 1993) have concluded that the efficiency of SSCP is highly sequence dependent and, thus, may vary significantly among fragments, depending on running conditions optimized for each locus. Alterations in single-strand DNA mobility in nondenaturing electrophoretic gels may be predicted by DNA folding computer software (Nielsen et al. 1995), but for most
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


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.