Every DNA sequencing project involves two steps: (i) making suitable templates for all the regions to be sequenced; and (ii) running sequencing reactions and electrophoresis. The latter step can be automated by use of workstations and autosequencers. The former step requires careful experimental design and laborious DNA manipulations such as the construction of nested deletion mutants (1). This is often the limiting step in large sequencing projects. The ‘shot-gun’ method eliminates this complicated DNA manipulations (1), but many recombinant clones must be sequenced because of the random nature of this procedure. Here we describe a novel sequencing technique that utilizes recent advances in amplification of long DNA fragments by PCR (2). This systematic method requires minimal amount of starting DNA and eliminates complicated steps for template preparation. We have successfully used this method to determine the sequence of a cosmid insert and the genomic structure of the transforming growth factor β type II receptor gene (5).

In our protocol, nested DNA fragments around region of interest are amplified by anchored PCR using a vectorette unit and the fragments are directly sequenced. The procedures are schematically presented in Figure 1.

Amplification primers were designed in the sequenced regions of the cosmid using the MacVector program (Kodak). The following oligonucleotides, V-top 5'-GAAGGAGAGGACGCTGTCTGTCGAAGGTAAGGAACGGACGAGAGAAGGGAGAG-3' and V-bottom 5'-CTCTCCCTTCTCGAATCGTAACCGTTCGTAACGAAATCGCTGTCCTCTCCCTC-3' were synthesized, purified and dissolved in distilled water to a final concentration of 4 µM of each. The solution was heated at 68°C for 10 min and was slowly cooled over 30 min to room temperature to make an annealed vectorette unit (4 µM; ref. 3). V-top and V-bottom are complementary to each other except for the middle one-third, giving the vectorette unit a bubble-like structure (Fig. 1). M13 sequence-tagged 224 primer (224M13: 5'-TGTAAAACGACG9GCCAGTCACTCGGTAACCGGAGGAGAAGGGAGAG-3') (3) was phosphorylated in a 50 µl volume containing 1× kinase buffer [70 mM Tris–HCl (pH 7.6), 10 mM MgCl₂, 5 mM dithiothreitol, 1 mM ATP], 50 µM 224M13 and 50 U T4 polynucleotide kinase. The reaction was incubated at 37°C for 30 min, heated to 68°C for 10 min then stored at −20°C.

Cosmid DNA was extracted from a 1.5 ml overnight culture by an alkaline mini-prep method (4) into 50 µl distilled water. Two microliters of this cosmid DNA solution were enzymatically digested using AluI, BsaAI, BstUI, PalII, Rsal, AccI, AflIII, BstYI, HinclI, MsiI, Tsp45I, EcoRV, HpaI, PvuII, Scal, SmaI, SspI or StuI (Stratagene and New England Biolabs) in a 20 µl volume including 10 U enzyme and 1× buffer as recommended by the manufacturers. After a 1 h incubation at 37°C, individual digestions were extracted with phenol–chloroform and precipitated with ethanol. Because AccI, AflIII, BstYI or Tsp45I do not give blunt-end DNA fragments, samples digested with them were blunt-ended by treatment with T4 DNA polymerase in a 50 µl reaction containing 1x buffer [50 mM NaCl, 10 mM Tris–HCl (pH

---

* To whom correspondence should be addressed
The PCR reactions were performed in 100 µl volumes containing 1× XL buffer II (Perkin-Elmer), 1.1 mM Mg(OAc)₂, 200 µM dNTPs, 2.5 µl from the vectorette unit-ligated restriction fragment solution, 1 µl 30 µM phosphorylated 224M13, 1 µl 30 µM amplification primer and 2 U rTth XL DNA polymerase. Reactions were heated to 94°C for 1 min, then PCRed for a total of 40 cycles at 94°C for 30 s, 55°C for 30 s, 68°C for 4 min; the last 24 cycles require a 15 s extension per cycle using a thermal cycler (Perkin-Elmer).

Amplified PCR fragments were purified in 50 µl distilled water using the Wizard PCR prep kit (Promega). Purified PCR fragments, 25 µl, were subjected to λ exonuclease digestion using the ‘PCR template prep for ssDNA sequencing’ kit (Pharmacia) to get single-stranded DNA in 25 µl of TE (10 mM Tris–HCl pH 7.6, 1 mM EDTA). Seven microliters of single-stranded DNA solutions were subjected to direct fluorescent sequencing using the Taq dye-primer cycle sequencing kit and the 373A DNA sequencer (Perkin-Elmer).

We used restriction enzymes with various lengths of recognition sequences (4–6 bases) to obtain various lengths of the restricted fragment around the amplification primer. This, in turn, enables us to generate various lengths of amplified fragments. As expected, the sizes of the amplified fragments were distributed up to 6 kb in length, giving a set of nested fragments suitable for long range sequence determination (Fig. 2a). The M13 sequence is at the end of the 224M13 primer, enabling sequencing from that end using a commercially available dye-primer sequencing kit. The sequences were usually quite readable (Fig. 2b), allowing long sequences to be read in a single run. After these sequences were determined and assembled new amplification primers were designed, synthesized and the steps beginning from PCR amplification were repeated using the same vectorette-ligated restriction fragment solutions.

Using this method, we could readily determine the sequence of 14 out of the 16 kb region of interest from a single small scale cosmID DNA preparation. We found, in total, a region of 2 kb with few restriction sites for the initial set of enzymes which required the use of additional restriction enzymes.

In addition to sequencing long continuous stretches of DNA, this strategy is suitable for nucleotide sequence determination around a region of known sequence. For example, we have successfully applied this method to determine the genomic structure of the transforming growth factor β type II receptor gene using YAC and cosmID clones (5).

The vectorette unit was originally used to isolate the end fragments from yeast artificial chromosome (YAC) clones because of its high specificity and low background. Using computer-designed amplification primers, we rarely encountered false-priming products.

DNA fragments amplified by the anchored PCR have been used to obtain nucleotide sequences adjacent to known sequences in the genome (6). Our method is characterized by the extensive use of vectorette-mediated anchored PCR by which a series of nested DNA fragments suitable for sequencing can be obtained in a single step. The ease of this technique, the use of minimal amounts of DNA and the ability to systematically sequence large nucleotide segments make this method advantageous and preferable to existing methods in various projects. This method alone or in conjunction with other methods should accelerate a wide variety of sequencing projects.

ACKNOWLEDGEMENTS

We thank M. McMenamin and E. Spillare for critically reviewing the manuscript.

REFERENCES

5 Takemoshita,S., Hagiwara,K., Nagashima,M., Gemma,A., Bennett,W.P. and Harris,C.C. Genomics, submitted.