0273-2289/95/5303—0195$04.60

Accelerated Article

A Strategy to Sequence Repetitive DNA Based on Partial Restriction Enzyme Cleavage

F. G. C. ABATH*,[1] AND A. A. HOLDER[2]

[1]*Centro de Pesquisas Aggeu Magalhaes-FIOCRUZ, Cidade Universitaria, CEP: 50670-420, Recife, Brazil; and* [2]*National Institute for Medical Research, Mill Hill, NW7 1AA, London, UK*

Received February 23, 1995; Accepted March 7, 1995

ABSTRACT

The strategy to sequence repetitive DNA described in this article is based on partial restriction enzyme cleavage. It is an alternative to using nested deletion with exonuclease III or similar enzymes in which progressively more remote regions of the target DNA are brought into range for sequencing by universal primers.

Index Entries: Sequencing; repetitive DNA; partial endonuclease cleavage; PCR.

INTRODUCTION

During a sequencing project in which a gene for a major tegumental antigen of *Schistosoma mansoni* was cloned and characterized (*1*), a strategy was devised to sequence on both strands a 964 bp repeat sequence within the open reading frame. The region of repeats posed some technical problems as it was too long to be sequenced using the dideoxy chain termination technique (*4*) from flanking primers and determine the point of overlap. There are claims that modern automated fluorescence-based DNA sequencers can separate and identify 800–1000 bases in a sequencing reaction (*2*). However, this equipment is very expensive and beyond the reach of most laboratories. The strategy described in the present paper is an alternative to using nested deletion with exonuclease III or

*Author to whom all correspondence and reprint requests should be addressed.

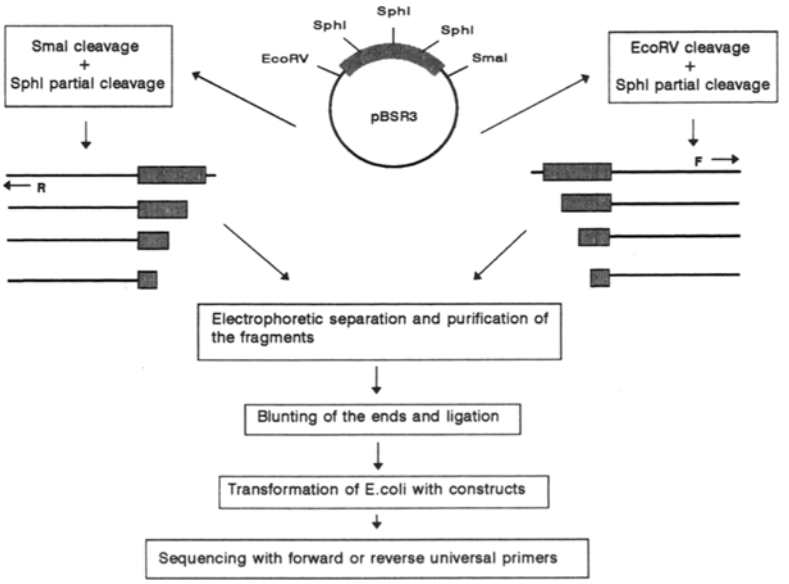

Fig. 1. Diagram depicting a strategy to sequence repetitive DNA subcloned in pBluescript. *See* text for a detailed description. F, universal primer M13-20 and R, universal primer M13 reverse.

similar enzmes in which progressively more remote regions of the target DNA are brought into range for sequencing by universal primers (*3*).

First, the size of this region was estimated by polymerase chain reaction (PCR), as described (*1*), using the primers GA70FP3 (5′-GCTTAATG-AACGCAGGTG-3′) and GA70RP0 (5′-CGAAGGTGAAGTGTACTC-3′), at the boundaries of the repetitive sequence. The number of repeats was determined using a primer complementary to one extremity of the region GA70FP3 and A70RT (5′-GATAAACTAGGCACTGGATCGT—3′), a primer complementary to the repeated elements. The first PCR amplified a fragment of approx 1 kb, the second PCR produced fragments of approx 960, 750, 453, and 220 bp (results in agreement with the final sequencing data [*1*]). Having estimated the size of the region of repeats, sequencing using flanking primers and long electrophoretic runs (in our hands 500 bases could be determined from each single primer) produced an overlapping sequence from each end. A strategy to sequence the repeats on both strands was devised based upon the presence of 3 *Sph*I sites in this region (Fig. 1). The DNA was subcloned in pBluescript II KS+ (Stratagene) and the recombinant phagemid (clone BSR3) was then cleaved close to the 5′ or 3′ end of the insert with *Eco*RV or *Sma*I respectively, generating blunt ends. Partial cleavage allowed the removal of different length *Sph*I fragments. *Sph*I partial digests were resolved in 1% agarose gels, and the purified bands were treated with T4 DNA polymerase to remove pro-

truding 3′ termini (*3*). The different partial digests were religated, the products used to transform *Escherichia coli* XL1-Blue and the constructs were sequenced using forward or reverse universal primers, M13-20 (5′-GTAAAACGACGGCCAGT-3′) and M13 reverse (5′-AACAGCTATGAC-CATG-3′), respectively.

The 964 bp repeat region comprised five major repeated elements. The second, third, and fourth were identical, 249 bp in length and containing a *SpHI* site. The first repeated element was 180 bp long and the fifth repeated element was truncated (for details about this region *see* ref. *1*).

We have used this simple approach to solve a specific problem related to the sequencing of the repeats in a gene encoding a major tegumental antigen of *Schistosoma mansoni.* However, the technique can be readily applied to other genes, the only criterion being that there are suitable restriction sites within the repeats and the resulting fragments can be revolved by electrophoresis.

ACKNOWLEDGMENTS

This work was supported by the British Medical Research Council (MRC), the Brazilian Council for Scientific Development (CNPq), and FACEPE. We are grateful to Paul Hagan for critical reading of the manuscript.

REFERENCES

1. Abath, F. G. C., Hagan, P., Jeffs, S. A., Schechter, I., Meadows, H. M., Holder, A. A., and Simpson, A. J. G. (1993), *Mol. Biochem. Parasitol.* **60,** 81.
2. Grothues, D., Voss, H ., Stegemann, J., Wiemann, S., Sensen, C., and Zimmermann, J. (1993), *Nucleic Acids Res.* **21,** 6042.
3. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989), *Molecular Cloning: A Laboratory Manual,* 2nd ed., Cold Spring Harbor Laboratory, New York.
4. Sanger, F., Nicklen, S., and Coulson, A. R. (1977), *Proc. Natl. Acad. Sci. USA* **74,** 5463.