PROGRESS REPORT Grant DE-FG02-90-ER-610

A Probe-Based Mapping Strategy for DNA Sequencing with Mobile Primers

L. D. Strausbaugh and C. M. Berg

The specific areas of research targeted for the period of this Progress Report included three general phases: 1. Optimization of probe-mapping by both the development of new transposons and the design of stream-lined methods for mapping; 2. Application of transposon-based methods to larger plasmids and cosmids; and 3. Initiation of PCR-based applications of transposons. Some aspects of the work described in this Progress Report have been conducted in collaboration with D. M. Berg (DOE Grant DE-FG02-89ER-60862 to D. Berg and C. Berg).

An important activity conducted in parallel with the research activities of the past year was the dissemination of the methods we have developed. The adoption of transposon-based approaches is the final indication of the success of this research. We have developed a user-friendly approach to the description of transposon-based methods and have aggressively promoted this technology in several large forums. Presentations have been made at the DOE Workshop, the Cold Spring Harbor Meeting on Genome Mapping and Sequencing, and the American Human Genetics Association meetings. In addition, we have written a review for Methods in Enzymology to be included in the volume on recombinant DNA methods. We are presently considering the submission of an application to the National Center for Human Genome Research to conduct a course focused on transposon-based methods of genomic analysis. Finally, we are involved in a number of collaborations with research groups at our home institution and elsewhere (Perkin-Elmer Cetus, Gibco BRL, BIOS Corp., Pharmacia LKB, Milligen, Pfizer, Dupont, Stratagene, Vanderbilt University, the University of Texas, Yale University, University of Medicine and Dentistry of New Jersey). Each of these collaborations facilitates and promotes the widespread adoption of transposon-based approaches.

1. Optimization of Probe-mapping.

Currently, probe-mapping remains the single most powerful and cost-effective method for the localization of transposons in cloned DNA fragments of all sizes. The optimization of probe-mapping is the primary goal of this project and encompasses several research initiatives, including the design of maximally efficient mapping strategies, the development and characterization of engineered derivatives of the transposon gamma delta, and the continued successful extension of our approaches to increasing numbers of cloned DNA fragments.

In our initial demonstrations of probe-mapping (Strausbaugh et al., 1990), we determined both the localization and orientation of transposons inserted in cloned fragments. This dual determination requires two hybridization steps and is necessary to demonstrate the random nature of the geometry of gamma delta insertions. While we continue to perform two sets hybridization in the experiments described herein to verify the randomness of orientation (as well as to determine location) of gamma delta insertions, only the single hybridization step that reveals the location of the transposon on the linear map of the cloned fragment is necessary for virtually all applications. For example, for sequencing or the synthesis of probes from transposon primer binding sites, it is desirable to use reactions in both directions from the transposon insertion; data are collected from both transposon-borne primers and sequence information is arranged by knowledge of the order of the insertions. This modification of probe-mapping significantly reduces the labor involved and renders the procedure much more compatible with non-radioactive detection methods since it is unnecessary to remove probe for a second round of hybridizations. In addition to the use of radiolabeled probes, we have used a commercially available non radioactive labeling and detection methods (ECL, Amersham and Photo-Gene, BRL), these procedures do not significantly improve the

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

efficiency of probe-mapping compared to radioactive methods. However, these non-radioactive methods do offer the substantial advantages that the probes are stable over long periods of time, thereby eliminating the repeated labeling reactions required with radiolabeled probes.

For the numerous reasons described in our initial application, we continue to concentrate on the development of the transposon gamma delta. One of the goals for the past year was the generation of small derivatives appropriate for multiplex sequencing and probe-mapping. Starting with mini-gamma delta-1 (a 1.8 kb element constructed from the res and kan gene of Tn5 and the 38 bp repeats of gamma delta), a series of PLEX derivatives have been generated. Each of five mini-gamma delta PLEX transposons contains a different pair of unique primers and a different pair of unique tags:

my δ -pLex transposons (1.9 kb)

These PLEX derivatives have been designed for immediate use in pre-existing libraries. Probe-mapping is easily accomplished with a single enzyme for inserts cloned into compatible sites for any of the following: Not I, BamHI, ClaI, EcoRI, HindIII, BglII, PstI, or SmaI, as long as there are no internal recognition sites for the enzyme chosen for the analysis in the cloned fragments. This collection of enzymes includes many historically used in library construction and we have been able to easily devise simple single-enzyme mapping schemes for all recombinant plasmids analyzed to date. It has been possible to devise simple schemes even if enzymes other than those listed above are used in cloning.

Using mini-gamma delta-1 or its PLEX derivatives, transposon insertions have been obtained and mapped in a variety of cloned molecules:

<u>Plasmid (size)</u>	<u>Vector</u>	<u>Features</u>
pDmHJ6.7 (13.7kb)	pBR325 (ECORI site)	D. melanogaster DNA; regions of highly heterogenesis base composition
pVE1322 (9.3 kb)	pUC19 (Bam HI)	S. a'ermittilis DNA; high G & C base composition
p 4.3 (7.5pb)	pUC (ECORI site)	D. melanogaster DNA; unknown base composition
PBJ (17.4)	pUC (Sst site)	Chloroplast DNA; A & T rich

To date, the linear map of mutagenized molecules translate into essentially random distributions of insertion sites, although the transposon inserts predominantly into local AT-rich "valleys" (Strausbaugh et al., 1990; Zhang, Woycik, Strausbaugh and Berg, in preparation).

28

During the past year we have also initiated a collaboration with Dr. Mark Freeman of Pharmacia LKB Biotechnology to use mapped transposon insertions in automated sequencing. Transposon-borne primer binding sites are especially well-suited to automated methods since only one pair of fluorescent labeled primers will be synthesized for sequencing an entire cloned insert of over 4 kb. The mutagenesis and mapping have been completed and we await results from the automated sequencing.

2. Application of transposon-based methods to large fragments cloned in plasmids and cosmids.

Since mapped transposon insertions are powerful tools for sequencing, probe production, and subcloning of large cloned fragments, it is important to demonstrate success with large plasmids and cosmids. We have used mini-gamma delta-1 to mutagenize a plasmid of total size 17.4 kb, consisting of a 14.7 kb fragment of chloroplast DNA cloned in pUC. This recombinant DNA molecule has provided our first and, to date, only example of problems associated with transposon-based approaches. Approximately 40% of the recovered mutagenized molecules contain deletions of different sizes, although the unmutagenized plasmid shows no instability. The simplest interpretation of the mapping experiments is that the transposon insertions occur randomly and, in certain geometries deletion occurs to a fixed region in the cloned fragment. This interpretation is based on complicated mapping results and must be proven by sequencing. According to our collaborator on this project, Dr. Robert Jansen of the University of Texas, this cloned fragment includes a region of the chloroplast DNA known to be involved in genome rearrangement. We cannot presently eliminate the possibility that this chromosomal "hotspot", in combination with transposon biology, is the source of the deletions.

We have recently undertaken the first experiments to analyze cosmid clones using transposons. Dr. Robert Donnelly (UMDNJ) has provided us with a 50 kb cloned human DNA fragment that contains the gene for the human interferon receptor protein. Dr. Donnelly and Dr. Sidney Petska have devised a new approach for mapping cloned fragments to chromosomes; this strategy is dependent on their ability to obtain probes from discrete positions distributed evenly over the linear map of the cloned fragment. In this collaboration, we will provide them with mutagenized cosmids that will then be used to synthesize probes from the transposon-borne primer binding sites. This collaboration is a particularly attractive vehicle for the development of cosmid applications for transposon technologies since relatively large (compared to sequencing) distances between insertion sites are desirable. Although the technical limitation for probe-mapping cosmids is the level of resolution of current gel systems (the accuracy of mapping decreases with increasing sizes of the fragments to be resolved), we are guaranteed success at the level required for the particular application in this collaboration. We are excited about the prospect of interfacing transposon technologies with physical mapping and chromosome assignment. The first round of mutagenesis on the cosmid clone has been completed and we are analyzing the results. In addition, we have used specialized gel system (FIGE) to fractionate large DNA segments in anticipation of the cosmid work.

3. Initiation of PCR-based applications of transposons.

The use of transposon-borne primer binding sites for PCR is a powerful approach with numerous applications, assuming that high fidelity, long-distance PCR can be attained. With support from Perkin-Elmer Cetus, we are involved in a project to accurately measure and improve the fidelity of PCR. Our collaborator D. Berg is improving the range and reliability of PCR to long distances (6-10 kb).

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

END

**DATE
FILMED
418192**

I