4096 hexamers, more than adequate for high throughput DNA sequencing by primer walking.

DOE Grant No. DE-AC02-76CH00016.

References

(1) Kieleczawa, J., Dunn, J. J., and Studier, F. W. DNA sequencing by primer walking with strings of contiguous hexamers. Science, 258, 1787-1791 (1992).

(2) Dunn, J. J., Butler-Loffredo, L. and Studier, F. W. Ligation of hexamers on hexamer templates to produce primers for cycle sequencing or the polymerase chain reaction. Anal.Biochem. 228, 91-100 (1995).

Rapid Preparation of DNA for Automated Sequencing

John J. Dunn, Matthew Randesi, and **F. William Studier**
Biology Department; Brookhaven National Laboratory; Upton, NY 11973
516/344-3012, Fax: -3407, *dunn@genome1.bio.bnl.gov*
http://genome5.bio.bnl.gov

We have developed a vector, referred to as a fesmid, for making libraries of approximately 35-kbp DNAs for mapping and sequencing. The high efficiency lambda packaging system is used to generate libraries of clones. These clones are propagated at very low copy number under control of the replication and partitioning functions of the F factor, which helps to stabilize potentially toxic clones. A P1 lytic replicon under control of the lac repressor allows amplification simply by adding IPTG. The cloned DNA fragment is flanked by packaging signals for bacteriophage T7, and infection with an appropriate T7 mutant packages the cloned sequence into T7 phage particles, leaving most of the vector sequence behind. The size of the vector portion is such that genomic fragments packageable in lambda (normal capacity 48.5 kbp) should also be packaged in T7 (normal capacity 40 kbp).

We have made fesmid libraries of several bacterial DNAs, including Borrelia burgdorferi (the cause of Lyme disease), Bartonella henselae (the cause of cat scratch fever), E. coli, B.subtilis, H. influenzae, and S. pneumoniae, some of which have been reported to be difficult to clone in cosmid vectors. Human DNA is also readily cloned in these vectors. Brief amplification followed by infection with a gene 3 and 17.5 double mutant of T7, which is defective in replicating its own DNA, produces lysates in which essentially all of the phage particles contain the cloned DNA fragment. Simple techniques yield high-quality DNA from these phage particles. Primers for direct sequencing from the ends of fesmid clones have been made.

Primer walking from the ends of fesmid clones could be an efficient way to sequence bacterial genomes, YACs, or other large DNAs without the need for prior mapping of clones. The ends of fesmids from a random library provide multiple sites to initiate primer walking. Merging of the elongating sequences from different clones will simultaneously generate the sequence of the original DNA and determine the order of the clones. The packaged fesmid DNAs are a convenient size for multiple restriction analyses to confirm the accuracy of the nucleotide sequence.

DOE Grant No. DE-AC02-76CH00016.

A PAC/BAC End-Sequence Database for Human Genomic Sequencing

Glen A. Evans, Dave Burbee, Chris Davies, Trey Fondon, Tammy Oliver, Terry Franklin, Lisa Hahner, Shane Probst, and Harold R. (Skip) Garner
Genome Science and Technology Center and McDermott Center for Human Growth and Development; University of Texas Southwestern Medical Center at Dallas; Dallas, TX 75235-8591
214/648-1660, Fax: -1666, *gevans@swmed.edu*
http://mcdermott.swmed.edu

While current plans call for completing the human genome sequence in 2003, major obstacles remain in achieving the speed and efficiency necessary to complete the task of mapping and sequencing. As an approach to this problem, we proposed a novel approach to large scale construction of sequence-ready physical clone maps of the human genome utilizing end-specific sequence sampling. An earlier pilot project was initially carried out to develop a GSS (genomic sequence sampled) map of human chromosome 11 by sequencing the ends of 17,952 chromosome 11 specific cosmids. This chromosome 11-specific end-sequence database allows rapid and sensitive detection of clone overlaps for chromosome 11-sequencing.

In this project, we propose to evaluate the utility of PAC and BAC end-sequences representing the entire human genome as a tool for complete, high accuracy mapping and sequencing. In this approach, we utilized total genomic PAC/BAC libraries (constructed by P. de Jong, RPCI), followed by end-sequencing of both ends of each clone in the library and limited regional mapping of a subset of clones as sequencing nucleation points by FISH (Fluorescence in situ hybridization).

To initiate regional analysis, a single clone would be sequenced by shotgun or primer directed sequencing, the entire sequence used to search the end-database for overlapping clones, and the minimal overlapping clones for extending the sequence selected. This approach would allow rational and efficient simultaneous mapping and sequencing, as well as expediting the coordination and exchange of information between large and small groups participating in the human genome project.