Abstract

Bacterial artificial chromosome (BAC) libraries, BAC-based physical maps, and whole genome sequencing maps have been under development for many genomes of the most important eukaryotic multicellular species, including Arabidopsis, Drosophila, rice, mouse, and human. However, DNA libraries built with conventional vectors and methods are biased, and genome gaps exist in all of the physical and sequencing maps. To optimize BAC cloning, physical mapping, and genome sequencing we have developed an optimized BAC cloning system. The CopyRight™ pSMART BAC 2.0 consists of the pSMART-BAC 2.0 vectors and the BAC-optimized Replicator® 2.0 competent cells. This BAC vector lacks an indicator gene and associated promoter, has termination signals on either side of the insertion site, and simplifies BAC preparation by selectively using single- or multi-copy number replicons. The vector shows much higher stability of inserts containing AT-rich sequences, direct and inverted repeats, and other deleterious DNAs, thus making it possible to construct unbiased BAC libraries. The CopyRights pSMART BAC 2.0 vector and the Replicator 2.0 cells also feature inducible amplification of copy number, increasing yields to as many as 50 copies per cell. The amplification is more robust than the similar CopyControl system and permits easy isolation of BAC DNA for sequencing, subcloning, or restriction mapping. To overcome the other gaps introduced by partial restriction digestion, we have successfully developed techniques to construct unbiased, randomly-sheared BAC libraries with large inserts (>100 kb). With these novel techniques and tools, we have cloned previously unclonable DNA and are seeking toward covering the entire genomes of many plants and animal species. We offer the Random Shear BAC Libraries as a custom service.

Background

Current BAC Libraries are constructed from partial digestions of genomic DNA. However, despite using multiple libraries, many gaps remain in all genomes studied (below). These gaps include, but are not limited to, repetitive DNA and genomic DNA was randomly sheared, size-selected, and cloned into the pSMART BAC vector. A 5X coverage library was screened with overgo oligonucleotide probes specific for various regions of Chromosome 1. Significantly, clone coverage across all the probed regions, including the centromeric region, was similar in the random shear library (Figure 4). In contrast, these regions show vastly different representation in the Arabidopsis genome project (15, 75, or <1 clone per 0.1 Mb, respectively; 17X coverage overall). Most importantly, we have been able to close existing centromeric gaps of this “fragmented” physical and sequence genomic map. The same probes also identified clones covering centromeric regions of other chromosomes.

Random Shearing of Genomic DNA

Mega-base regions of genomic DNA, such as centromeres, may completely lack recognition sites for common restriction enzymes (e.g., BamHI, EcoRI, HindIII; Figure 2, left).

Lucigen has developed methods to randomly shear genomic DNA into fragments of 100-400 kb. Significantly, the DNA from all genomic regions is sheared into fragments of ~4-6 kb are often deleted when cloned into standard BAC vectors (Figure 6).

Reduced bias in pSMART BAC vector

Teharyzome genomic DNA (75% AT) is very difficult to clone. Fragments as small as 4-6 kb are often deleted when cloned into standard E.coli vectors. However, unlike other BAC vectors, this vector does not induce high-level expression of insert DNA, further increasing stability of recombinant clones.

Summary

Improved vector for BAC libraries. A transfection-free BAC vector provides more stable cloning with very low background.

BAC Optimized competent cells. Electromediated transfection of the new pSMART BAC vector incorporates CloneSmart transcription-free technology to increase the stability of cloned inserts. In addition, a unique system selects against non-recombinant clones (Figure 5). However, unlike other BAC vectors, this vector does NOT induce high-level expression of insert DNA, further increasing stability of recombinant clones.

Random shearing for reduced bias. Random sheared libraries are a powerful tool for closing genomic gaps, including centromeric regions.

We have successfully contracted random shear BAC libraries from 8 plants and 5 animal species, including the important models: Medicago truncatula, Xenopus tropicalis, and mouse. More importantly, we have constructed random shear BAC libraries of genomic DNA from crustacean species, which is impossible to digest by restriction enzymes.

BAC libraries. Lucigen offers custom BAC library using standard partial restriction digestion or random shearing. These techniques provide unparalleled performance for obtaining large, unbiased BAC libraries.