Whole Genome Shotgun Library Approach For Microbial Sequencing Projects at the JGI

Eileen M. Dalin, Doug Smith, Hope Tice, Kerrie Barry, Alla Lapidus, David Bruce, and Paul M. Richardson

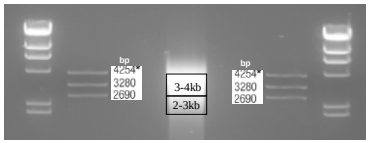
DOE JOINT GENOME INSTITUTE
US DEPARTMENT OF ENERGY
OFFICE OF SCIENCE

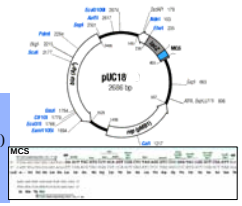
The US Department of Energy's Joint Genome Institute is a high-throughput sequencing center and user facility that has sequenced a large number of microbial genomes. The strategy for most projects calls for construction of whole genome shotgun libraries from high-molecular weight DNA isolated from an axenic culture. In general, the JGI produces 3 insert size-selected libraries for all whole genome shotgun projects. We generate a 3kb high-copy pUC18 library, an 8kb low-copy pMCL200 library, and a 40kb pCC1FOS fosmid library. The DNA is randomly sheared, fragments are end-repaired for blunt-end cloning, and then size selected on an agarose gel, extracted and purified. 3 & 8kb inserts are cloned into the appropriate vector and transformed into *E. coli.* 40kb inserts are cloned, packaged and infected by phage into *E. coli.* PCR using primers flanking the inserts are used to determine the percentage of clones with inserts for both the 3 and 8kb libraries, before proceeding to production sequencing. Clones (10-384-well plates) from each of the 3 & 8kb libraries are initially sequenced and library quality is assessed at this stage before full sequencing is completed. Both 3 & 8kb libraries are sequenced to 4x sequencing coverage and the 40kb library is sequenced to 30x clone coverage. The 3 library approach generally results in more complete genome coverage at the draft stage, and pairing information allows for contig mapping and repeat resolution during the genome assembly/finishing step.

LBNL-60350 Poster

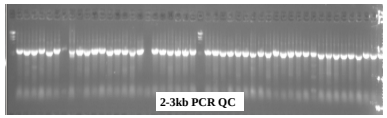
Small Insert: 3kb Library Construction

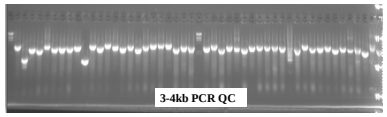
- Randomly shear 3-5ug of genomic DNA
- Blunt-end Repair
- Size select on gel
- Extract 2 band sizes of 2-3 & 3-4kb
- Gel purify
- Blunt-end ligate into pUC18
- Transform into electromax DH10B cells
- PCR QC 24 clones of each library

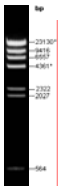




- **Name:** pUC18 (2.7kb)
- **Replicon:** pMB1
- **Purpose:** High Copy (~500/cell)
- **Selectable Marker:** Amp
- **Color Selection:** lacZ
- **Cloning Site:** SmaI

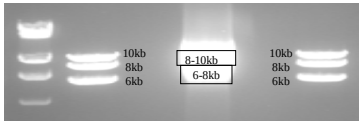


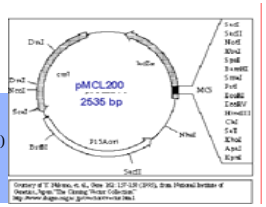




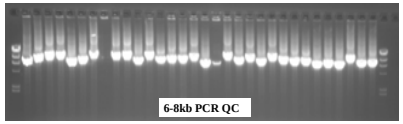
Mid-size Insert: 8kb Library Construction

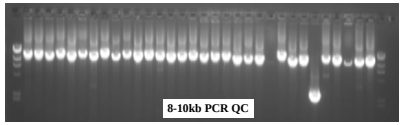
- Randomly shear 10-20ug of genomic DNA
- Size select on gel
- Extract 6-10kb fragment
- Gel purify
- Blunt-end Repair
- Size select on gel
- Extract 6-8, & 8-10kb fragments
- Gel purify
- Blunt-end ligate into pMCL200
- Transform into electromax DH10B cells
- PCR QC 24 clones of each library





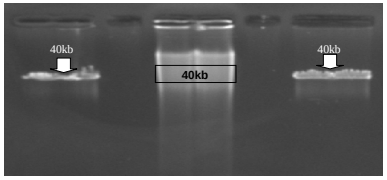
- **Name:** pMCL200 (2.5kb)
- **Replicon:** p15A
- **Purpose:** Low Copy (10 copies/cell)
- **Selectable Marker:** cml
- **Color Selection:** lacZ
- **Cloning Site:** EcoRV

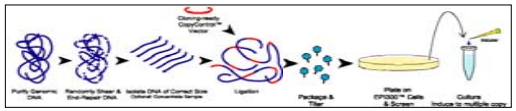


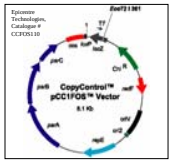


Large Insert: 40kb Library Construction

- Randomly shear 20ug of genomic DNA
- Size select on a pulse-field gel
- Extract 40kb fragment
- Gel purify
- Blunt-end Repair
- DNA Cleanup
- Blunt-end ligate into pCC1FOS
- Package into phage
- Infect *E. coli*

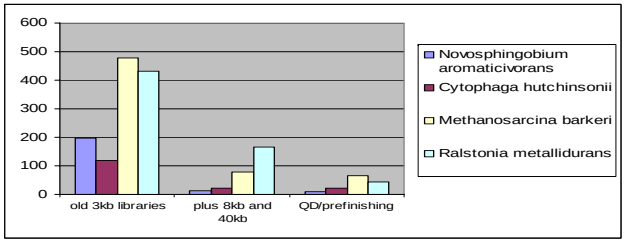






Microbial Assembly/Finishing

	old 3kb lib.		plus 8kb and 40kb		QD/prefinishing	
	Major Contigs	Genome size (MB)	Major Contigs	Genome size (MB)	Major Contigs	Genome size (MB)
Novosphingobium aromaticivorans	**197**	4.17	**13**	4.21	**9**	4.215
Cytophaga hutchinsonii	**118**	4.36	**23**	4.41	**22**	4.41
Methanosarcina barkeri	**478**	3.88	**77**	4.83	**67**	4.84
Ralstonia metallidurans	**432**	NA	**165**	6.83	**45**	6.83



This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231 and Los Alamos National Laboratory under contract No. W-7405-ENG-36.