Library Construction using Trace Amount of DNA

Julianna Chow, Eileen Dalin, Tanja Woyke, Susan Lucas, and Jan-Fang Cheng

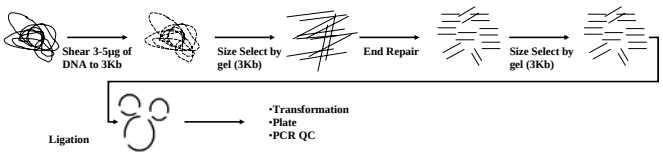
DOE JOINT GENOME INSTITUTE
US DEPARTMENT OF ENERGY
OFFICE OF SCIENCE

Abstract

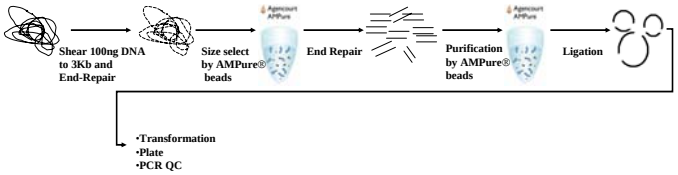
As a user facility, the US Department of Energy's Joint Genome Institute, in collaboration with scientists around the world, are able to generate DNA sequences for a diversity of organisms and environmental samples. Often times, the amount of DNA provided for library construction is limited. It is important to develop a protocol that requires a trace amount of DNA for library construction. In an attempt to test the minimum amount of DNA necessary for library construction, we decided to try a different approach, using AMPure® beads instead of agarose gels, to purify DNA. AMPure® bead purification, utilizes solid-phase paramagnetic bead technology to purify DNA fragments from contaminates and enzymes, with minimal loss of DNA. We performed the test on a chloroplast genome of a flowering plant Brighamia insignis because of its small genome size (~130Kb) and the availability of a finished genome. The DNA concentrations used in this experiment ranges from 100 to 1,000 ng. By using AMPure® beads, we eliminate the need for gel separation to purify and size select DNA fragments, which requires 3-5 ug of starting material. The AMPure® bead purified DNA are cloned into the pUC18 vector and sequenced to assess the quality of the libraries. The sequencing data from the libraries constructed using AMPure® bead are compared with the library constructed using the standard gel purification method to determine the cloning efficiency, insert size distribution, coverage biases, and assembly accuracy.

Comparing 3Kb Library Construction Approaches

Standard 3Kb Library Construction



Library Construction with AMPure Beads



Shearing Low Concentration of DNA

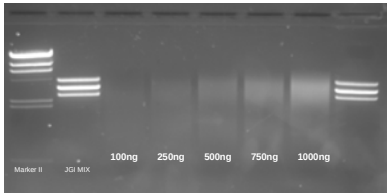


Figure 1. This experiment is to understand the shearing condition of low concentration genomic DNA. Different quantity of genomic DNA were sheared using a HydroShear. Lambda DNA ranging from 1 to 10 ng/ul was sheared using the same setting and separated on 1% agarose gel. Different concentrations of DNA generate similar sizes of fragments under the same shearing condition.

Library Construction Using AMPure® Beads

A.

Library Name	Starting material (ng)	DNA remaining after 1st purification (ng)	% of DNA loss	DNA remaining after 2nd purification (ng)	Loss between 1st and 2nd purification	Total % Loss
FXHH	100	32.80	67.21%	22.20	32.31%	77.80%
FXHN	250	79.10	68.36%	52.40	33.75%	79.04%
FXHO	500	183.75	63.25%	142.40	22.50%	71.52%
FXHP	750	211.75	71.77%	156.40	26.14%	79.15%
FXHS	1000	340.90	65.91%	238.00	30.18%	76.20%

B.

Library name	Starting material (ng)	Colony Count	Single transformation (CFUs)	# of possible 384's at 60%
FXHH	100	25	5000	7.81
FXHN	250	127	25400	39.69
FXHO	500	640	128000	200.00
FXHP	750	498	99600	155.63
FXHS	1000	1460	292000	456.25

C.

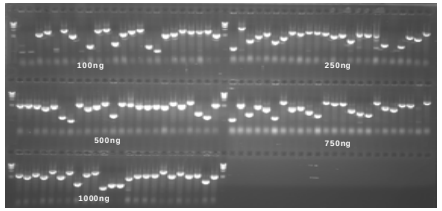


Figure 2. Libraries of different starting material were sheared to 3Kb using a hydroshear. After shearing, the samples were purified using AMPure® beads and end repaired. The samples were purified using AMPure® beads for a second time to remove enzyme and contaminates from the end-repair reaction. The samples were ligated with a 2:1 (insert:vector) ratio using Fast-Link Ligase kit. Unfortunately, there was a greater than 75% loss of DNA after two purifications, as indicated (A). Following a 1.5 hour ligation, the samples were transformed in to Electromax DH10B cells. There is a decline in colony forming units (CFUs) as the amount of starting material were reduced (B). The quality of the library were evaluate by PCR and running a QC PCR agarose gel (C). Based on the QC agarose gel, majority of the libraries contain DNA inserts size of about 3Kb.

Insert Size Distribution

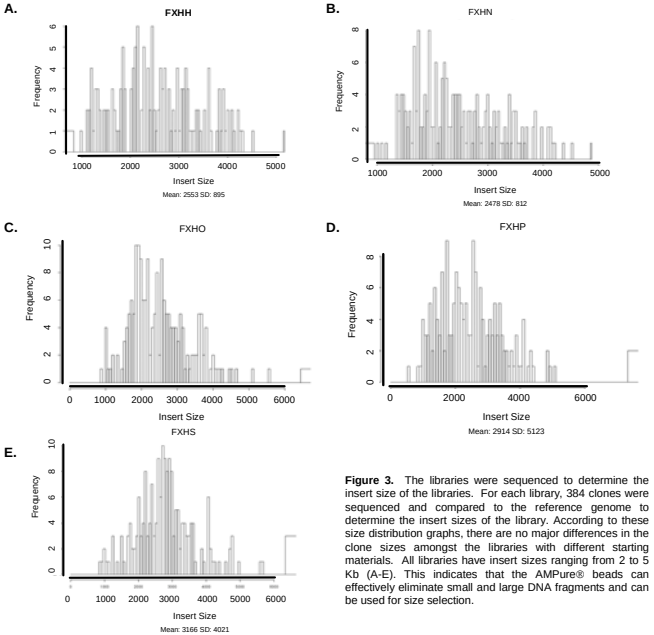


Figure 3. The libraries were sequenced to determine the insert size of the libraries. For each library, 384 clones were sequenced and compared to the reference genome to determine the insert sizes of the library. According to these size distribution graphs, there are no major differences in the clone sizes amongst the libraries with different starting materials. All libraries have insert sizes ranging from 2 to 5 Kb (A-E). This indicates that the AMPure® beads can effectively eliminate small and large DNA fragments and can be used for size selection.

Comparison of Libraries Constructed With Different Purification Methods

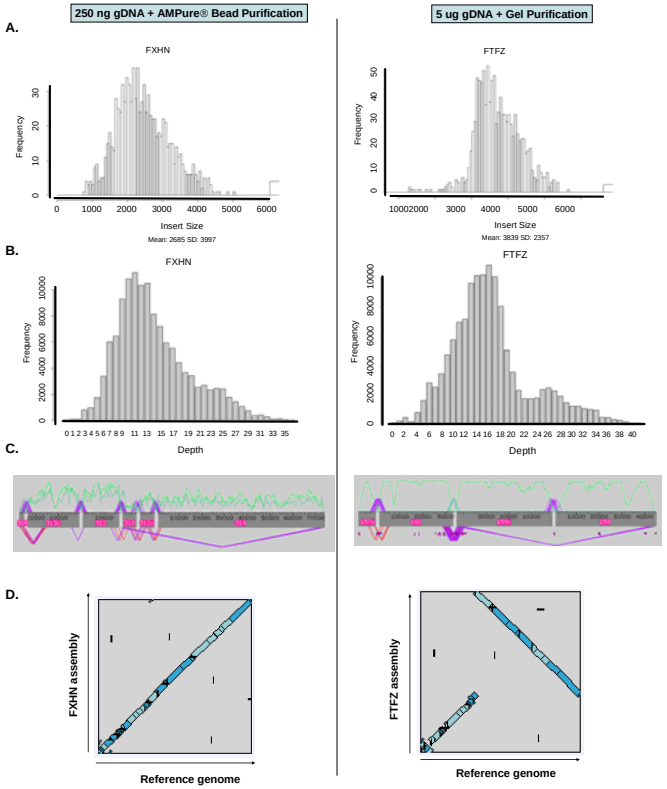


Figure 4. FXHN is a library constructed using AMPure® beads with 250 ng of chloroplast DNA (left panel) and FTFZ was constructed using gel purification with 5 ug of the same chloroplast DNA (right panel). For each library, 3,840 reads were generated and compared to the reference genome to determine the insert size distribution (A), the coverage of the genome (B), the number of gaps in the assembled sequences (C), and the accuracy of the assembled sequences (D). It appears that the library created using the AMPure® beads has a higher standard deviation in the clone sizes. The genome coverage of the two libraries seem to have no significant differences. The assembly of the library created using the gel purification method has a couple of mis-assemblies in the repeat regions of the chloroplast genome.

Conclusions

1. **AMPure® beads allow the possibility of constructing Sanger libraries with trace amount of starting material.**
2. **Cloning efficiency remains high even when small amount of DNA inserts are used and can generate enough sequencing reads to assemble Metagenomes and Prokaryotes genomes.**
3. **Libraries constructed using the AMPure® beads have a larger size distribution than the gel purification method.**
4. **Libraries created with AMPure® beads assemble as well as libraries created using the standard gel purification.**

Acknowledgements

We would like to thank Mansi Choviata for colony picking; Alicia Clum, Penny Yi Pang and David Robinson for sequencing; Michael Zhang for developing in house software to analyze data; and Matt Nolan for assembly.

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 Berkeley National Laboratory under contract No. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396.