# Ligation-mediated PCR Amplification as a Tool to Finish Microbial Genomes

**Hope N. Tice**, Eileen M. Dalin, Ze Peng, Susan Lucas, and Jan-Fang Cheng

DOE JOINT GENOME INSTITUTE
US DEPARTMENT OF ENERGY
OFFICE OF SCIENCE

## Abstract

As Sanger sequencing is being replaced by higher throughput and lower cost of next generation sequencing, finishing microbial genomes will face two major challenges. First, the technology will need to be fast enough to handle many more drafted genomes. Second, it will have to incorporate a clone-free approach to fill gaps. We have been testing a method that utilizes a universal "bubble-tag" to perform primer walking and gap closure in a clone-free condition. The "bubble-tag" method was first described by Doug Smith (PCR Methods Appl. 2: 21-27, 1992) to amplify and sequence lambda DNA. There is no evidence however, that this approach will work for the more complex microbial genome. Here we describe the experimentation of this approach in primer walking of the E. coli genome. Genomic DNA was sheared, blunt-end repaired, and ligated to the bubble adaptors. Site specific primers were used together with the universal bubble primer to amplify the regions of interest. We applied the Ampure beads binding and washing step to reduce the amount of small amplified fragments. The remaining large amplified DNA appears to be suitable for sequencing. Different bead-to-DNA ratios were tested in order to generate long amplified templates. This approach enables primer walking and gap filling in a clone-free sequencing process. More importantly, the uniformity of this approach is amenable for an automated finishing process.

## AMPure Beads Effective Removal of Small Fragments

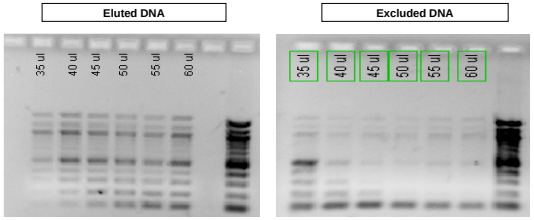



**Figure 1. A test was set up using a DNA ladder to determine the effectiveness of AMPure beads in the removal of small fragments. The volume of DNA used was 50µl in each reaction. The volume of beads used is listed above the wells with the control loaded in the last well on the right. The ratios of beads-to-DNA ranged from 0.7 to 1.2.**

## Bubble PCR Primer Walk Workflow

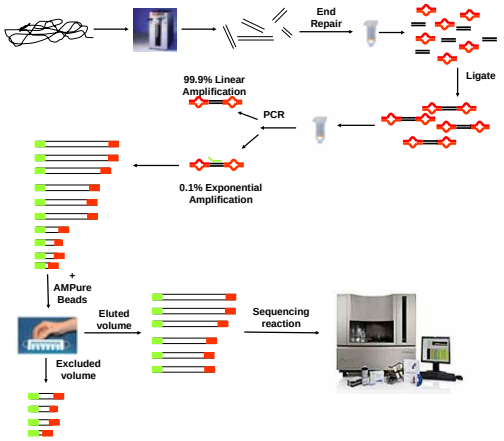



**Figure 2. The DNA of interest is randomly sheared to 3kb fragments. After end repair the DNA is cleaned up by a column. Bubble adapters are then ligated to the DNA ends. Another cleanup is done and PCR reactions are set up. The reactions contain the bubble primers and the site-specific primers with M13 tails. In our test experiment, about 99.9% of the fragments would have a linear amplification whereas 0.1% would have an exponential amplification. The PCR reaction is cleaned up using AMPure beads to remove small fragments. Once the desired sample is eluted from the beads, the sequencing reaction is done.**

## The Test of Different Ampure Bead to DNA Ratios

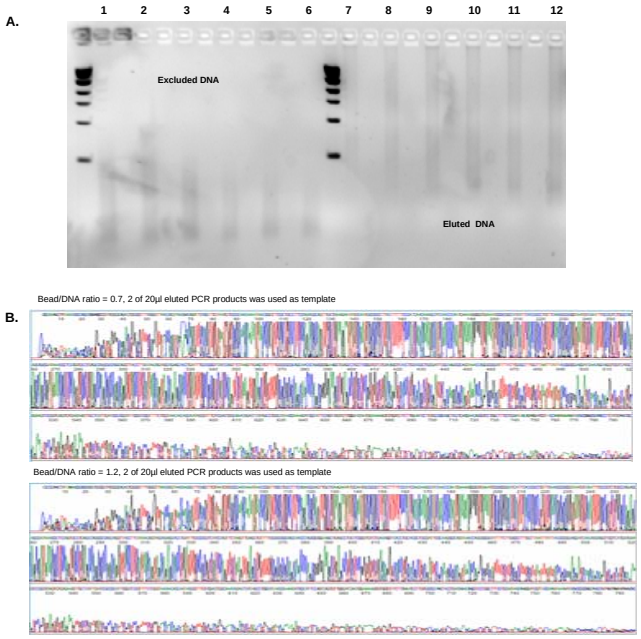



**Figure 3. The beads-to-DNA ratios of 0.7 to 2.0 were used to see which one provided the most effective cleaning of small fragments. (A) PCR amplified products were cleaned by AMPure beads using the following beads-to-DNA ratios: 0.7 (lanes 1 and 7), 0.8 (lanes 2 and 8), 0.9 (lanes 3 and 9), 1.0 (lanes 4 and 10), 1.1 (lanes 5 and 11), and 1.2 (lanes 6 and 12). The chromatograms generated from sequencing templates of lanes 7 and 12 are shown in (B). The beads-to-DNA ratios in this range for cleaning the small fragments do not seem to significantly affect the quality of the sequencing reads.**

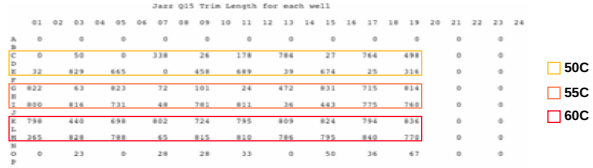
## E. Coli Sequences Used in the Primer Walk

| Oligo Name | Sequence (5' to 3') | lengh | Start | End | Strand |
|---|---|---|---|---|---|
| Exo-FW | GTTTTCCCAGTCACGACGTTGTA**CGTGGCTGTTCATTTGCTTA** | 43 | 3022954 | 3022973 | + |
| Exo-RV | GTTTTCCCAGTCACGACGTTGTA**TCTGGTAAACGCTGATGACG** | 43 | 3023148 | 3023129 | - |
| OmpA-FW | GTTTTCCCAGTCACGACGTTGTA**TGACCGAAACGGTAGGAAAC** | 43 | 1072627 | 1072646 | + |
| OmpA-RV | GTTTTCCCAGTCACGACGTTGTA**TGAGTACGCGATCACTCCTG** | 43 | 1072761 | 1072742 | - |
| GlyS-FW | GTTTTCCCAGTCACGACGTTGTA**ATACAGACGCACCGCTTCTT** | 43 | 3818465 | 3818484 | + |
| GlyS-RV | GTTTTCCCAGTCACGACGTTGTA**TGAATGAGCAGTATCAGCGG** | 43 | 3818709 | 3818690 | - |
| TopA-FW | GTTTTCCCAGTCACGACGTTGTA**TCGATCATTTCGACCATTCA** | 43 | 1419982 | 1419981 | + |
| TopA-RV | GTTTTCCCAGTCACGACGTTGTA**CCCTCTCTCCGGATCTTTTTC** | 43 | 1420221 | 1420202 | - |
| HolA-FW | GTTTTCCCAGTCACGACGTTGTA**CTGCCATACCCGATGCTTAT** | 43 | 609221 | 609240 | + |
| HolA-RV | GTTTTCCCAGTCACGACGTTGTA**CAGCCAACCGGTTATTTTGT** | 43 | 609441 | 609422 | - |
| CadA-FW | GTTTTCCCAGTCACGACGTTGTA**GTGAGTGGACTGGGTTTCGT** | 43 | 4455905 | 4455924 | + |
| CadA-RV | GTTTTCCCAGTCACGACGTTGTA**CAGCACGCTACCATTGCTAA** | 43 | 4456162 | 4456143 | - |
| ExuR-FW | GTTTTCCCAGTCACGACGTTGTA**AAAAATGTGGACCCAGCGTAG** | 43 | 3343811 | 3343831 | + |
| ExuR-RV | GTTTTCCCAGTCACGACGTTGTA**AATAGATAGCGGTCGGCATTG** | 43 | 3343142 | 3343123 | - |
| PcnB-FW | GTTTTCCCAGTCACGACGTTGTA**TAGCCCGCTTGTAGCAGTTT** | 43 | 132458 | 132477 | + |
| PcnB-RV | GTTTTCCCAGTCACGACGTTGTA**GCGTTATCCGTCTGATTGGT** | 43 | 132662 | 132643 | - |
| EutC-FW | GTTTTCCCAGTCACGACGTTGTA**AATACGTTTCGCCAAATCCAC** | 43 | 2646237 | 2646256 | + |
| EutC-RV | GTTTTCCCAGTCACGACGTTGTA**CGCGTGAAGATTGAAGATCA** | 43 | 2646469 | 2646450 | - |
| Nth-FW | GTTTTCCCAGTCACGACGTTGTA**CGCCTTTTGAATTGCTGATT** | 43 | 1800200 | 1800219 | + |
| Nth-RV | GTTTTCCCAGTCACGACGTTGTA**ACGTTGGCTGTTTTACGACC** | 43 | 1800488 | 1800469 | - |

**Figure 4. Twenty primers were selected from the E. coli genome for the test of primer walk using the bubble-PCR approach. A M13 tag was added to each primer (in bold fonts) so that the PCR products can be used directly for sequencing. See figure 6.**

## 19 of 20 Primers Generated Sequences

**A. Different annealing temperatures in PCR can affect sequencing performance**




**B. Sequences are localized to where the primers are designed**

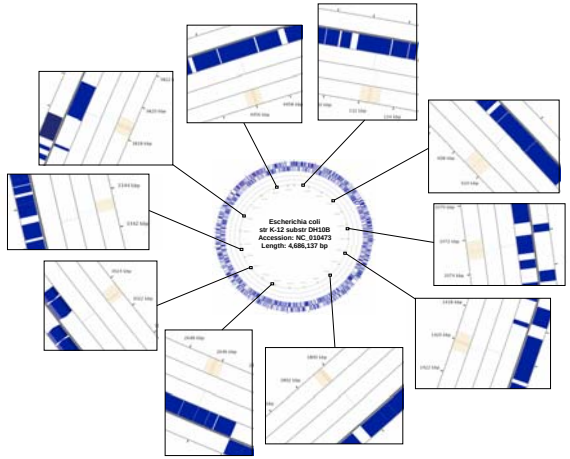



**Figure 5. (A) 19 of the 20 test primers successfully generated high quality sequences at 60C annealing temperature. (B) The produced sequences (in light brown color) are mapped to the 3' flanks of the designed primers (in light blue color) in the E. coli genome . This figure was generated using the on-line CGView Server.**

## Proposed Gap Closure & Primer Walking Process

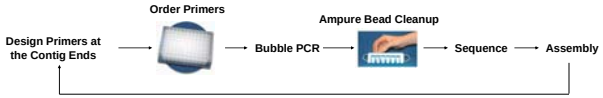



**Figure 6. The gap closure and primer walking processes in finishing can become a scalable process if automation is introduced. In our current sequencing pipeline, PCR set up, bead cleaning, and chemistry sequencing steps are streamlined already. With some informatics support at primer design and ordering, the proposed automated finishing process would become reality.**

## Conclusions

1. **Randomly sheared gDNA can be used in bubble PCR to generate templates for the finishing process**
2. **The 19 of 20 E. coli primers tested in this experiment successfully generate high quality sequences**
3. **This approach enables a clone-free finishing process that is amenable for automation**

## Acknowledgements

**We would like to thank Tanja Woyke for providing the E. coli primer sequences, David Robinson for helps in sequencing, and Dorothy Lang for aligning the sequences to the E. coli genome.**

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.