Gap Closing/Finishing by Targeted Genomic Region Enrichment and Sequencing

Kanwar Singh*, Jeff Froula*, Hope Tice, Len A. Pennacchio and Feng Chen

DOE JOINT GENOME INSTITUTE
US DEPARTMENT OF ENERGY
OFFICE OF SCIENCE

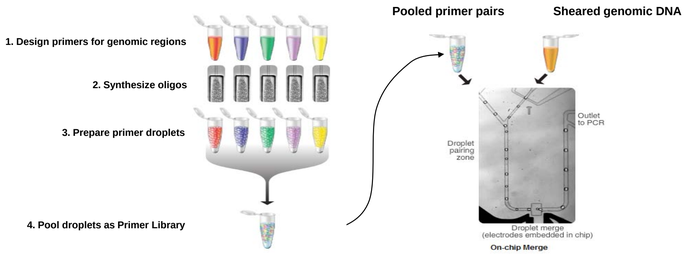
ABSTRACT

Gap Closing/Finishing of draft genome assemblies is a labor and cost intensive process where several rounds of repetitious amplification and sequencing are required. Here we demonstrate a high throughput procedure where custom primers flanking gaps in draft genomes are designed. Primer libraries containing up to 4,000 unique pairs in independent droplets are merged with a fragmented genomic template. From this millions of picoliter scale droplets are formed, each one being the functional equivalent of an individual PCR reaction. The PCR products are concatenated and sequenced by Illumina which is then assembled and used for gap closure. Here we present an overall experimental strategy, primer design algorithm and initial results.

OVERVIEW OF RAINSTORM MICRODROPLET BASED TECHNOLOGY FROM RAINDANCE

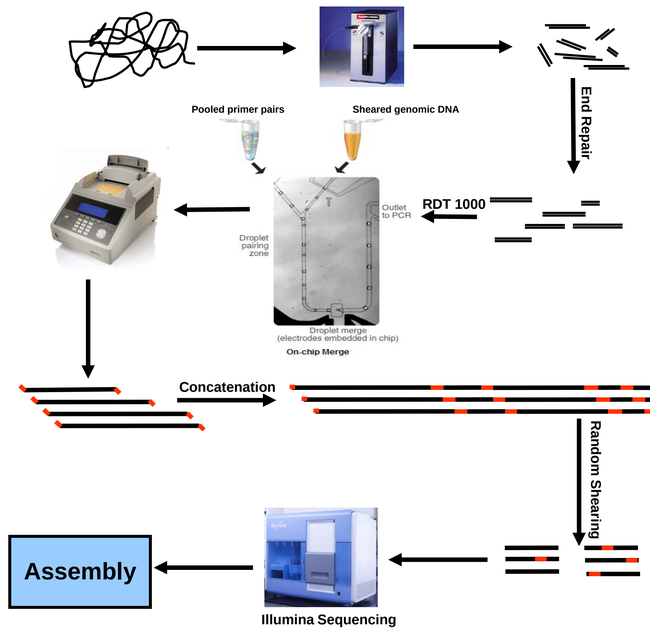
RDT 1000

Disposable chip



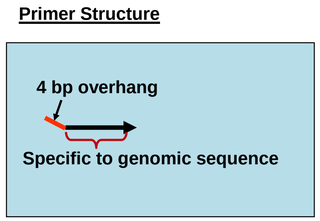
- Picoliter-volume droplet functions as an individual test tube
- Droplet can contain single molecule, cell, or bead
- 10 million droplet per hour
- No need for automated liquid handling
- Chip based and chips are disposable
- **Consistency**: 1% variation in droplet sizes
- **Stability:** can be thermocycled, shipped and stored
- **Control:** generation, combination, detection, & sorting

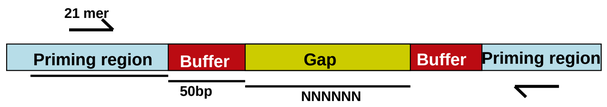
EXPERIMENTAL WORKFLOW



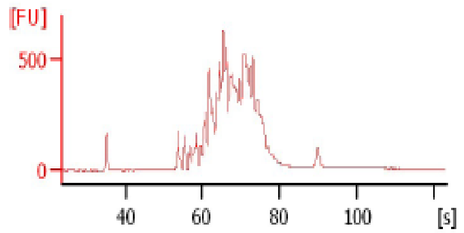
PRIMER DESIGN

In our initial study we assembled ***Pichia stipitis*** (15.4Mbp) with an in house assembler (Meraculous) resulting in 192 gaps. Primer design is described in the figure below.

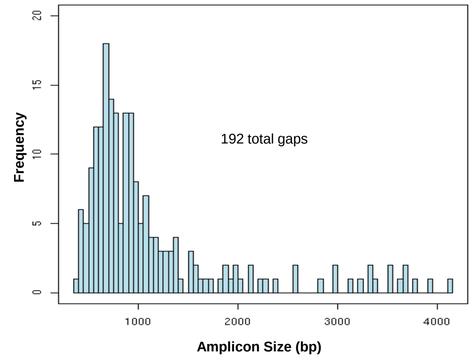




Bioanalyzer Profile of PCR Product

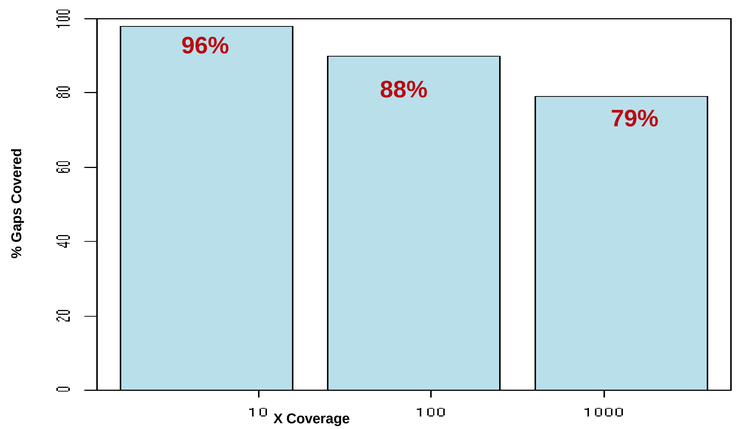


Histogram of Amplicon Sizes



RESULTS

Internal Gap Closure



Note: Out of the total 192 gaps, 1 failed completely and 4 where covered less than 10X. (I.e. at least one was in repeat region, others maybe PCR or assembly failure). The average coverage in gaps was ~ 2500X.

FUTURE EXPERIMENTS

After the first successful gap closing/finishing experiment, JGI recently acquired their own RDT 1000. Current experiments will try to
1) design and synthesize pooled primer libraries around the gaps of pooled genomes, thereby closing/finishing gaps in dozens of genomes in a single experiment
2) also, close scaffold gaps by designing primers using bubble PCR approach.

ACKNOWLEDGEMENTS

This work is performed in collaboration with RainDance Technologies. Primer pair library synthesis and droplet merging was performed at RainDance Technologies facility. We thank their participation in the study.

RainDance Technologies

* Equally Contributed to Work

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.