Highly accurate, long HiFi reads produced by the PacBio Sequel II system have brought new levels of contiguity, completeness, and accuracy to large genome assembly tasks. Here, we introduce a highly parallelizable shearing method, a streamlined library prep workflow, and an assembly algorithm based on HiFi reads that enable a high throughput, end-to-end solution for microbial genome assembly.

High raw accuracy workflow improvements workfolows

• Long reads improve assembly contiguity. However, once median read length exceeds the median repeat length of a genome, accuracy improvements yield greater gains in contiguity than longer reads.
• Because even very similar repeats can be distinguished with HiFi reads, shorter libraries can produce the same or better assembly results.
• Tolerance for shorter input DNA samples simplifies sample handling and allows closed genomes even from lower quality samples.
• Higher raw accuracy reduces the coverage requirement for assembly, increasing sample throughput and reducing cost per sample.

Scalable, low-cost DNA shearing

• The need for high molecular weight DNA can add complexity, cost, and time to long read library preparation. In addition, megaplot 3 system and g-TUBE shearing do not scale efficiently beyond 8 or 24 samples, respectively.
• We present a novel DNA shearing method compatible with HiFi sequencing using acid washed glass beads (Sigma Aldrich, G8772) and agitation on a Vortex Genie2.

Figure 2. Glass bead shearing of high molecular weight E. coli gDNA produces a high peak

• 300 ng of E. coli gDNA in 50 µL of water was added to 0.2–0.8 g of glass beads in 300 µL of water. Samples were vortexed for 15 min – 1 hr, 3,000 rpm.
• The distribution of fragment sizes can be adjusted by changing the mass of beads or the vortex time.

Figure 3. Glass bead shearing produces consistent shears even with variable quality input DNA

• The new SMRTbell library prep kit (SMK 3.0) has fewer steps, fewer reagents, and is automatable.
• Paired with glass bead shearing, SPK 3.0 creates an efficiently scalable high throughput workflow.

Figure 8. SMRTbell library prep workflow for up to 96 samples

• Input DNA QC (Qubit) 1 hour
• Glass bead shearing & cleanup 1.5 hours
• Repair & A-tailing 2 hour
• Adapter ligation & cleanup SPK 3.0
• Nucleotide treatment & cleanup 0.5 hour
• Library QC (Qubit) 1 hour
• Equal mass pooling & cleanup 30 min
• Final library QC & ionizing 1.5 hours

Figure 9. Microbes assembled with HiFi data have gold standard accuracy

• Expected plasmid sizes are noted above samples grouped by species. Plasmids in green were recovered in at least one replicate.
• Phased results of those using default parameter without curation. Manual review can improve contiguity (S. aureus USA300-TCH1618, E. coli H10407) and identify assembly artifacts.
• Recovery of smaller plasmids may require use of smaller insert sizes.

Conclusions

• Glass bead shearing is an inexpensive, robust method of producing DNA suitable for HiFi sequencing in a highly parallel manner, removing a significant barrier to high throughput SMRT sequencing of microbial genomes.
• The SPK 3.0 further simplifies the SMRTbell prep workflow, reducing handling times and enabling automation by eliminating size selection and time-consuming manual QC steps.
• The Microbial Assembly workflow in SMRT Link consistently produces highly accurate, closed microbial genomes from HiFi data with default settings, completing an end-to-end high throughput workflow for microbial whole genome sequencing.

References