Streamlined SMRTbell® library generation using addition-only, single tube strategy for all library types reduces time to results.

Jenny Ekholm[1], Jenny Gu[1], Jeff Burke[2], Michelle Kim[2], Justin Sibert[2], Kelvin Liu[2]
Pacific Biosciences, 1305 O'Brien Drive, Menlo Park, CA 94025, Circulomics Inc. 701 E Pratt St, Baltimore, MD 21202

ABSTRACT

We have streamlined the SMRTbell library generation protocols with improved workflows to deliver seamless end-to-end solutions from sample to analysis. A key improvement is the development of a single-tube reaction strategy that shortened hands-on time needed to generate each SMRTbell library, reduced time-consuming AMPure purification steps, and minimized sample-handling induced gDNA damage to improve the integrity of long-insert SMRTbell templates for sequencing. The improved protocols support all large-insert genomic libraries, multiplexed microbial genomes, and amplicon sequencing. These advances enable completion of library preparation in less than a day (approximately 4 hours) and opens opportunities for automated library preparation for large-scale projects.

Here we share data summarizing performance of the new SMRTbell Express Template Kit 2.0 representing our solutions for 10 kb and >50 kb large-insert genomic libraries, complete microbial genome assemblies, and high-throughput amplicon sequencing. The improved throughput of the Sequel System with read lengths up to 30 kb and high consensus accuracy (>99.999% accuracy) makes sequencing with high-quality results increasingly assessible to the community.

EXPRESS 2.0 WORKFLOWS

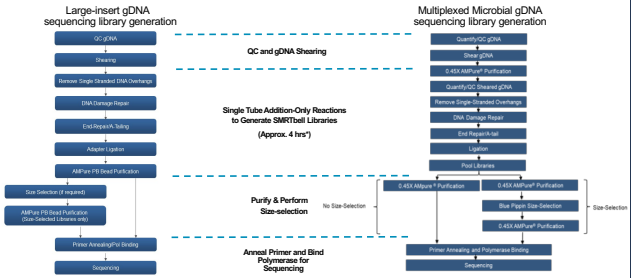

Figure 1. Detailed Express 2.0 workflows for long-insert and multiplexing microbial libraries

RECOMMENDED gDNA INPUT REQUIREMENTS

1a.

Insert Size Target	Size Selection	gDNA Quality Required	Shearing Method	Range of Sheared and Concentrated Input gDNA (µg)	Recommended Sheared and Concentrated DNA Amount
~10 kb	No	Mode >20 kb	g-TUBE	1.0 to 2.0 µg	1.0 µg
>15 kb (size-selected using BluePippin) system) >30 kb (size-selected using BluePippin system)	Yes	Mode > 40 kb	g-TUBE or Megaruptor®	2.0 to 5.0 µg	3.0 µg
	Yes	Mode >50 kb	26G needle or Megaruptor®	3.0 to 5.0 µg	5.0 µg

1b.

Insert Size Target	Size Selection	gDNA Quality Required	Shearing Method	Recommended microbial gDNA Amount
~12 kb	Optional	Predominantly >20 kb	g-TUBE	1.0 µg

Table 1a) and b). Recommended gDNA input requirement. (a) gDNA input requirements for different target sizes for long-insert libraries, (b) gDNA input requirements for microbial multiplexing libraries

EXPECTED SEQUENCING PERFORMANCE*

2a. Library Type with 3.0 sequencing chemistry

	Polymerase Read length (Average)	Polymerase Read Length (N50)	Longest subread (Average)	Longest subread (N50)
>30 kb Size-selected libraries	up to 30 kb	up to 45 kb	up to 35 kb	up to 40 kb
10 kb libraries with no size selection	up to 24 kb	up to 50 kb	up to 6.5 kb	up to 8.9 kb

2b. Library Type with 3.0 sequencing chemistry

	Polymerase Read length (Average)	Polymerase Read Length (N50)	Longest subread (Average)	Longest subread (N50)
~12 kb insert	up to 45 kb	up to 85 kb	up to 8 kb	up to 11 kb

Table 2a) and b). Expected sequencing performance. (a) for long-insert libraries of different target sizes, (b) for multiplexing microbial libraries

*Sequencing performance, reads/data per SMRT Cell and other expected results vary based on sample quality/type and insert size.

LARGE-INSERT gDNA QC AND SEQUENCING RESULTS

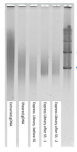

Figure 2. QC of >30 kb libraries. QC gel image for a >30 kb size-selected SMRTbell Express 2.0 library show library fragment smear extending up to 97 kb.

Large-insert gDNA library sequencing yield

Library Type	SMRTbell Library Yield	Sequencing Yield	SMRT Cells
SMRTbell insert libraries supported	Yield of SMRTbell libraries for sheared genomic DNA input.	Mapped bases per input ug of sheared genomic DNA	Minimum number of SMRT Cells supported by library
15 kb Size-selected Libraries	up to 27 %	up to 250 Gb	> 20
30 kb Size-selected Libraries	up to 20%	up to 35 Gb	> 4
10 kb Non-size-selected Libraries	up to 75%	up to 450 Gb	> 40

Table 3. Long-insert gDNA library sequencing yield for the different library target sizes

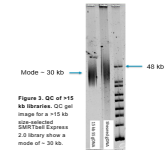

Figure 3. QC of >15 kb libraries. QC gel image for a >15 kb size-selected SMRTbell Express 2.0 library show a mode of ~ 30 kb.

MULTIPLEXED MICROBIAL gDNA QC AND SEQUENCING RESULTS

Experimental Design: 10-plex with a total of 42 MB pooled microbial genome size

Barcode ID	Sample ID	Gram Status	Genome Size (bp)
BC1001	Escherichia coli K12 MG1655 [1]	-	4,653,240
BC1002	Escherichia coli K12 [2]	-	4,653,240
BC1009	Klebsiella pneumoniae ATCC BAA-2146 [3]	-	5,781,501
BC1010	Bacillus cereus Strain 971 [3]	+	5,427,083
BC1012	Listeria monocytogenes CFSAN008105 [4]	+	3,032,269
BC1015	Shigella sonnei CFSAN030807 [4]	-	5,082,953
BC1016	Neisseria meningitidis FAM18 [3]	-	2,194,961
BC1018	Staphylococcus aureus subsp. aureus ATCC 25923 [3]	+	2,806,345
BC1019	E. coli Strain W [3]	-	5,005,347
BC1022	Staphylococcus aureus HPV107 [3]	+	2,901,406
		Total Genome Size	42,352,644

Table 4. Detailed information on the on the 10 microbial genomes tested

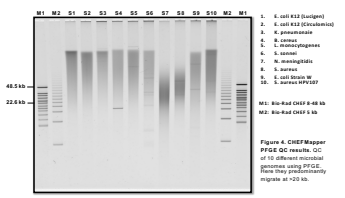

Figure 4. CHEF Mapper PFGE QC results. QC of 10 different microbial genomes using PFGE. Here they predominantly migrate at >20 kb.

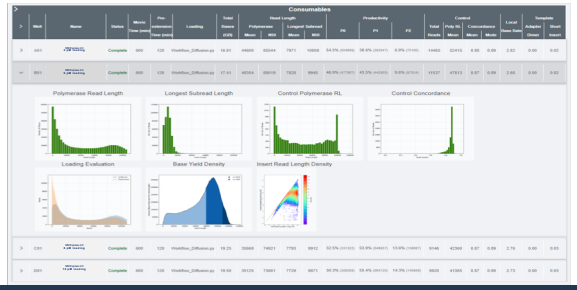

Figure 5. Sequencing results for pooled microbial genome libraries. Underloading and overloading can impact assembly results. We recommend an on-plate concentration of 6 to 12 pM adjusted to your experience with instrument for optimal results.

Multiplexed Microbial Genome Assembly Results

Barcode ID	Sample ID	Expected Genome Size (bp)	Genome Complexity Class	GC Content (%)	# Contigs (Mean chromosomal assembly)	Concordance w/ NCBI Reference (Q>)	Polished Contigs (#)	Max Contig Length (bp)	N50 Contig Length (bp)	Sum of Contig Lengths (bp)	Pre-Assembled Yield (%)	Filtered Subread Coverage (fold)	Assembly Notes
BC1001	*E. coli K12* (Lucigen)	4,653,240	I	50.8%	1	56	1	4,642,499	4,642,499	4,642,499	93.1%	64	Complete chromosomal assembly.
BC1002	*E. coli K12* (Circulomics)	4,653,240	I	50.8%	1	57	1	4,642,500	4,642,500	4,642,500	92.9%	55	Complete chromosomal assembly.
BC1009	*K. pneumoniae*	5,781,501	II	57.0%	1	48	5	5,435,746	5,435,746	5,746,850	92.5%	50	Complete chromosomal assembly, and 140 kb and 85 kb plasmid assemblies. 116 kb plasmid captured in 2 contigs. Missed 2 kb plasmid.
BC1010	*B. cereus*	5,427,083	I	35.3%	1	35	2	5,408,315	5,408,315	5,423,588	92.7%	59	Complete chromosomal and 16 kb plasmid assembly.
BC1012	*L. monocytogenes*	3,032,269	I	37.9%	1	46	2	3,043,149	3,043,149	3,137,529	93.7%	66	Complete chromosomal assembly.
BC1015	*S. sonnei*	5,082,953	II	51.0%	1	62	1	4,813,454	4,813,454	4,813,454	93.1%	53	Complete chromosomal assembly. Missing eight expected plasmids.
BC1016	*N. meningitidis*	2,194,961	I	51.6%	1	50	1	2,213,947	2,213,947	2,213,947	92.2%	74	Complete chromosomal assembly.
BC1018	*S. aureus*	2,806,345	I	32.9%	1	56	2	2,778,860	2,778,860	2,806,350	92.6%	92	Complete chromosomal and 27 kb plasmid assemblies.
BC1019	*E. coli Strain W*	5,005,347	III	50.8%	1	44	2	4,898,327	4,898,327	5,004,399	93.0%	68	Complete chromosomal and 103 kb plasmid assemblies. Missed 5 kb plasmid.
BC1022	*S. aureus HPV107*	2,901,406	III	32.9%	1	33	2	2,962,786	2,962,786	2,994,972	93.4%	82	Complete chromosomal and 24 kb plasmid assemblies.

Table 5. Multiplexed Microbial Genome Assembly Results. Summary of genome assembly results with respect to GC content and genomic complexity.

CONCLUSION

The SMRTbell Express Template Prep Kit 2.0 product family provides streamlined, expedited workflows for improved ease of use with better recovery and yield. This enables the completion of library generation within a day for sequencing.

For large-insert gDNA library, SMRTbell library construction can be completed in 4 hours.

For multiplexed microbial genome assemblies, the key improvement is the change in shearing condition to generate ~12 kb insert libraries, and the much simplified workflow to complete library preparation from gDNA to SMRTbell library in ~ 8.5 hrs, depending on number of samples being processed. Customers will appreciate the reduced AMPure purification steps required by the protocol, and the single-tube addition strategy opens up opportunities to explore automated library generation.

Forthcoming this year in Summer 2019, we anticipate new solutions for our full-length transcript Iso-Seq application and amplicon sequencing to join the Express TPK 2.0 product family.

REFERENCES

Procedure & Checklist - Preparing gDNA Libraries Using the SMRTbell® Express Template Preparation Kit 2.0

Procedure & Checklist - Preparing Multiplexed Microbial Libraries Using SMRTbell® Express Template Prep Kit 2.0

Analysis Procedure - Multiplexed Microbial Assembly with SMRT Link v6.0.0 and Express Template Prep Kit 2.0

Express Microbial Multiplexing Calculator