Analysis Procedure – Multiplexed Microbial Assembly with SMRT® Link v10.0 and SMRTbell® Express Template Prep Kit 2.0

For microbial multiplexing, use the following 16 barcoded adapters contained in the Barcoded Overhang Adapter Kits (8A and 8B):

- Barcoded Overhang Adapter Kit 8A [PN 101-628-400]: bc1001, bc1002, bc1003, bc1008, bc1009, bc1010, bc1011, bc1012
- Barcoded Overhang Adapter Kit 8B [PN 101-628-500]: bc1015, bc1016, bc1017, bc1018, bc1019, bc1020, bc1021, bc1022

Set Up a Run on a Sequel®, Sequel II or Sequel IIe System (Any Sequel System)

Open the Run Design module in SMRT Link. Click **New Run Design**. Fill in the Sample Information section, selecting Application **Microbial Assembly**. Click the small arrow to open **Barcoded Sample Options**. Specify the following options:

1. Sample is Barcoded: **Yes**
2. Barcode Set: **Sequel_16_barcodes_v3**
3. Same Barcodes on Both Ends of Sequence: **Yes**
4. Assign Bio Sample Names to Barcodes: Select barcodes used and assign bio sample names either interactively or from a file.
5. Save the Run Design and launch the sequencing run.

Review the Resulting Data Sets

Once the run is complete, open the SMRT Link Data Management module. The barcoded samples are automatically demultiplexed. Click the number in the Demultiplexed Subsets column (**16** in this example) to view the barcoded Data Sets in a separate table.

		Data Set Details >				Sample Details	
Name	Demultiplexed Subsets	Well Sample Name	Run ...	Date Created	Created By	Bio Sample Name	Barcode Name
Microbial_Multiplex...	16	Microbial_Multiplex_A...	Run 01.0...	2019-01-09, 03...	pbicsuser	[multiple]	[multiple]

From the Data Management module home page, click the Data Set name of the parent sample (**Microbial_Multiplex** in the above example) to view the Data Set details. The Barcodes section displays the number of reads per barcode and several graphs to assess the quality of the demultiplexing results. Refer to the PacBio SMRT Link User Guide (v10.0), Demultiplex Barcodes Application section, for further details on interpreting this report.

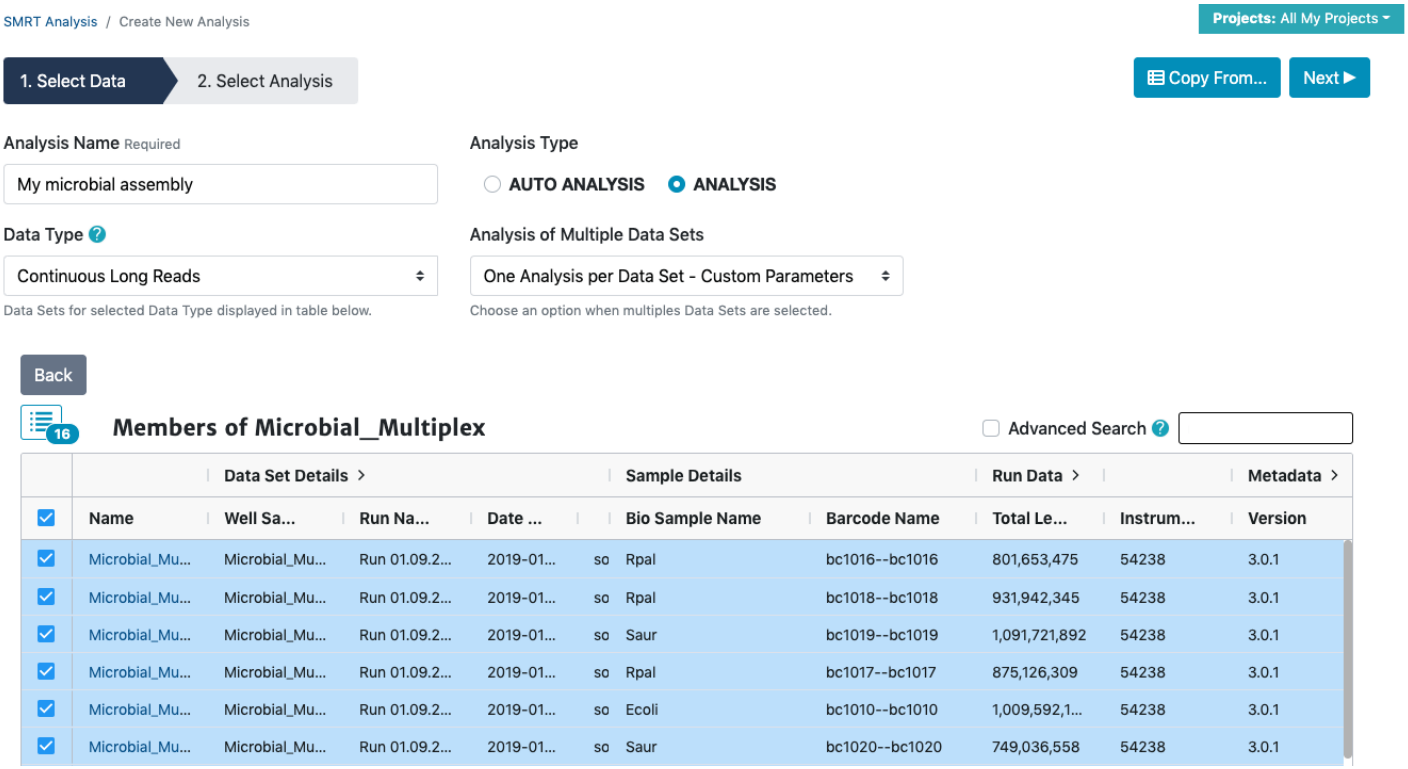
If review of the demultiplexing results reveals that any demultiplexing parameters need to be corrected (such as barcode set, barcode layout, or minimum barcode score), launch the **Demultiplex Barcodes** analysis application manually from the SMRT Analysis module, using the same Data Set as input.

Launch Microbial Assembly Analyses on the Microbial Data Sets

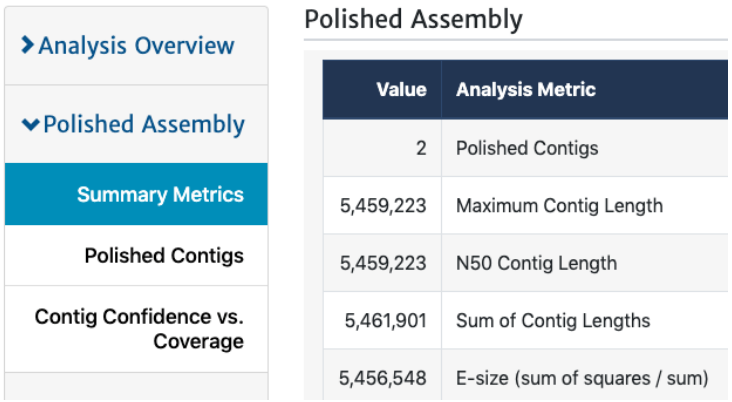
The Microbial Application in SMRT Link v10.0 provides *de novo* assembly, circularization, and plasmid identification for microbial genomes.

To launch one assembly analysis for each barcoded sample, open the SMRT Analysis module. Click **Create New Analysis**. Specify the following options:

1. Analysis Name: Enter an analysis name to persist across each analysis.
2. Data Type: Select **Continuous Long Reads**.
3. Data Sets: Locate the parent Data Set of the demultiplexed set (**Microbial_Multiplex**, in the above example). Click the number in the Demultiplexed Subsets column (16) to view the child data sets. Click the checkbox in the column header to select all children.
4. Analysis Type: **One Analysis per Data Set – Custom Parameters.**



5. Click **Next**.
6. Select Analysis Application: **Microbial Assembly.**
7. For the first Data Set, edit the **Genome Length** parameter if desired.
8. Click **Start and Create Next**. This launches the first assembly analysis on the first Data Set and proceeds to the launch page of the next Data Set, preserving any edits to the parameters made in the previous page.
9. Edit the **Genome Length** parameter as appropriate for the second Data Set.
10. Click **Start and Create Next**. Repeat the process for all Data Sets. After launching the last analysis, you will be redirected to the SMRT Analysis module home page, which displays the status of each launched analysis.
11. When an analysis completes, click its name to view the results.



Additional Troubleshooting Guidance

If the **Microbial Assembly** analysis generates a microbial genome assembly with >5 contigs, below are some additional suggestions to help improve the assembly, which can be tried individually or together.

1. To ensure sufficient coverage for assembly, set the **Genome Length** parameter to a value equal to (or slightly above) the actual genome size for the provided sample.
2. Increasing the coverage for assembly may help when the library quality is less than ideal.
   - In the **Create New Analysis** page, click **Advanced Parameters and** change the **Coverage** parameter from the default of 30 to a higher value (such as 40).
3. If the input subread length is relatively short (e.g. under 6000 bp) the minimum threshold for preassembled read length may need to be lowered. This can also help for longer insert libraries for which the quality is less than ideal.
   - In the **Create New Analysis** page, click **Advanced Parameters** and edit the **Advanced Assembly Options** by adding the following line, which modifies the overlap filtering options by specifying a smaller “--min-len" value of 2000 bp (default is 4000 bp):

     stage1.ovl_filter_opt = --max-diff 80 --max-cov 100 --min-cov 1 --bestn 20 --min-len 2000 --gapFilt --minDepth 4;
4. In case the Microbial Assembly analysis is taking a long time to complete, try modifying the block size.
   - In the **Create New Analysis** page, click **Advanced Parameters** and edit the **Advanced Assembly Options** by adding the following line, which changes the block size for the overlap process to a smaller value such as 200 Mb (default is 1024 Mb):

     stage1.block_size = 200; stage2.block_size = 200;

     Note that the block sizes for both stages do not have to be identical, but in general, the same rule of thumb applies for both stages.
5. In case the Microbial Assembly analysis stops due to an out-of-memory problem or overly high memory consumption in general, reducing the block size for the overlap process can help. Refer to the previous point for details on how to set the block size.
6. Optionally you can use the auto-analysis option when designing the sequencing run to set up analysis for automatic execution once the sequence data is available on the SMRT Link server. For more information, see the PacBio SMRT Link User Guide (v10.0).

Revision History (Description)	Version	Date
Initial release.	01	December 2020