# Technical overview: Multiplexed amplicon library preparation using SMRTbell prep kit 3.0

Sequel II and IIe systems ICS 11.0 / SMRT Link v11.1

PN 102-395-900 Version 03 (November 2022)

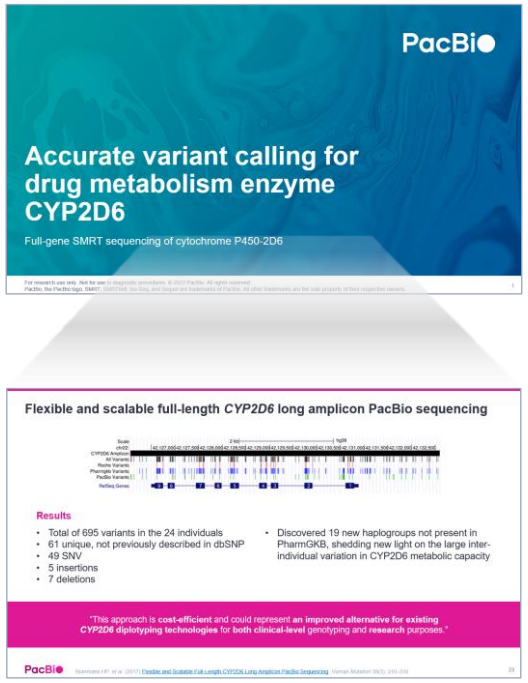
# Multiplexed amplicon library preparation using SMRTbell template prep kit 3.0

## Technical overview

1. Multiplexed amplicon library sample preparation workflow overview
2. Amplicon sample QC requirements
3. Multiplexed amplicon library preparation using barcoded primers or barcoded adapters
4. Multiplexed amplicon library preparation using barcoded M13 primers
5. Multiplexed amplicon library sequencing preparation workflow overview
6. Multiplexed amplicon data analysis recommendations
7. Sample preparation recommendations for full-length 16 amplicon sequencing
8. Sample preparation recommendations for human leukocyte antigen (HLA) amplicon sequencing
9. Sample preparation recommendations for HiFi target enrichment sequencing using Twist Bioscience panels
10. Technical documentation & applications support resources

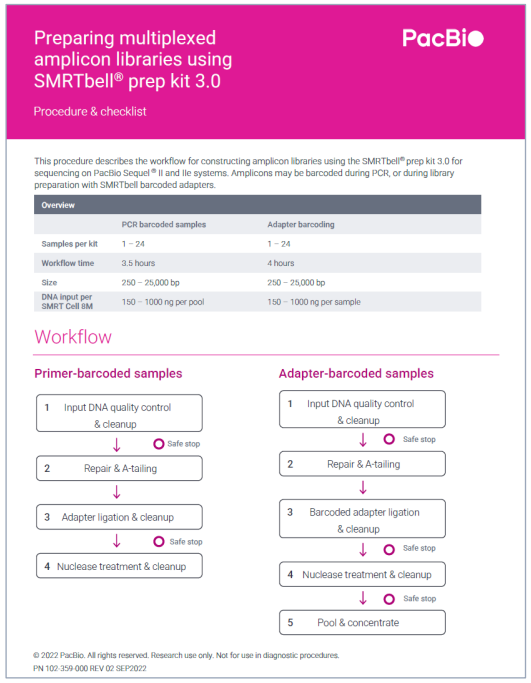
# Multiplexed amplicon sequencing: How to get started

**Application-specific educational literature** → **Application-specific Procedure & checklist** → **Application-specific technical overviews** → **Library construction, sequencing & analysis**



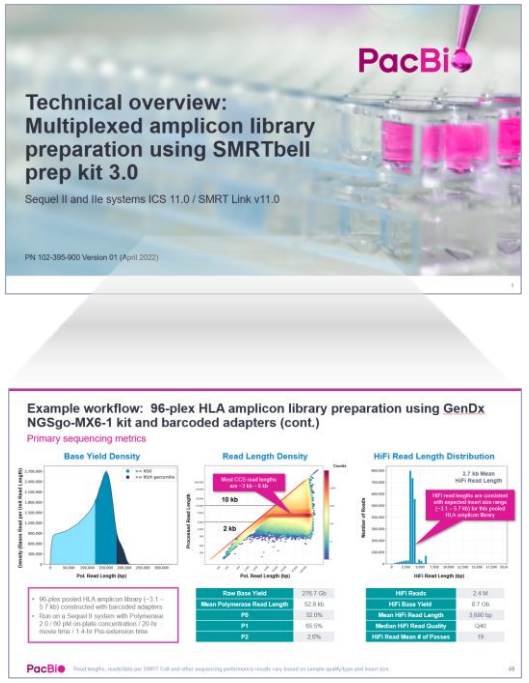

**PacBio literature website ( Link )**

Application-specific brochures, informational guides and and other product literature containing best practices recommendations for sample preparation and data analysis workflows.




**Procedure & checklist – Preparing multiplexed amplicon libraries using SMRTbell Prep Kit 3.0** (102-359-000) / **Procedure & checklist – Preparing multiplexed amplicon libraries using PacBio barcoded M13 primers and SMRTbell prep kit 3.0** (101-921-300)

Technical documentation containing sample library construction and sequencing preparation protocol details




**Technical overview: Multiplexed amplicon library preparation using SMRTbell prep kit 3.0** (102-395-900)

Technical overview presentations describe sample preparation details for constructing HiFi libraries for specific applications. Example sequencing performance data for a given application are also summarized.

## PCR amplification

Optionally use barcoded primers or barcoded M13 primers
Use ≥300 – 1000 ng of pooled barcoded amplicon input for library construction

## Library construction (SMRTbell prep kit 3.0)

Optionally barcode samples using barcoded adapters if starting with non-barcoded amplicons

## Sequencing (Sequel II and IIe systems)

ABC* with Sequel II binding kit:
Binding kit 3.1 (<3 kb amplicons)
Binding kit 3.2 (≥3 kb amplicons)
10 - 30-hr movie collection time

## Data analysis (SMRT Link or third-party)

* ABC = Anneal primer / Bind polymerase / Clean up bound complex

# Multiplexed amplicon library sample preparation workflow overview

# Multiplexed amplicon sequencing is supported with three barcoding options

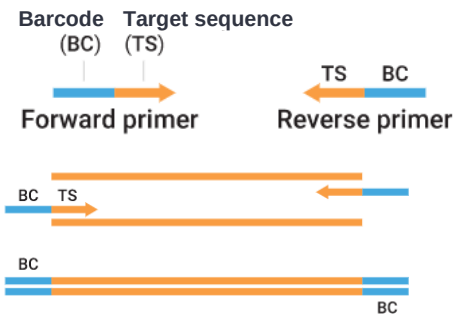
## Barcoded primers

**Use customer-supplied primers**

**Procedure & checklist – Preparing multiplexed amplicon libraries using SMRTbell Prep Kit 3.0** (102-359-000)

- Target-specific primers tailed with PacBio barcodes are used to produce symmetrically or asymmetrically barcoded samples with a 1-step PCR method.
- Barcoded amplicons are pooled into a single tube for SMRTbell library construction
- Can be used for small or large projects when validation of barcoded target-specific primers will be performed at the start of a project

### 1-Step PCR



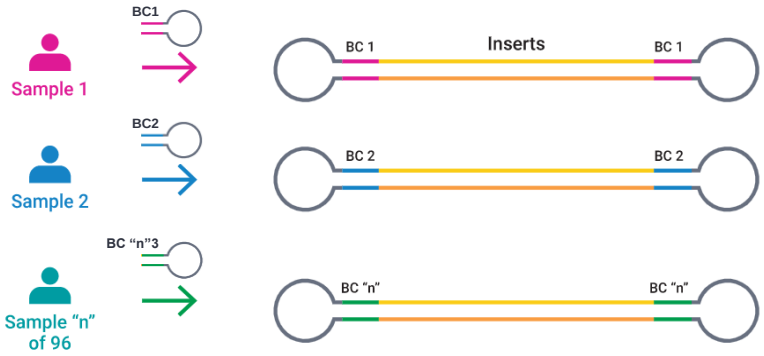

## Barcoded adapters

**PacBio SMRTbell barcoded adapter plate 3.0 (102-009-200)**

**Procedure & checklist – Preparing multiplexed amplicon libraries using SMRTbell Prep Kit 3.0** (102-359-000)

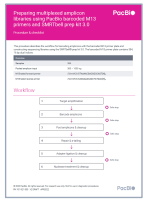
- PacBio barcodes are added to amplicons through ligation of barcoded adapters during SMRTbell library construction to produce symmetrically barcoded samples.
- Barcoded amplicons are pooled into a single tube for subsequent sequencing on a single SMRT Cell
- Recommended for smaller projects (≤96 samples) using validated PCR systems and off-the-shelf assays.

### SMRTbell barcoded adapter ligation




## Barcoded M13 primers

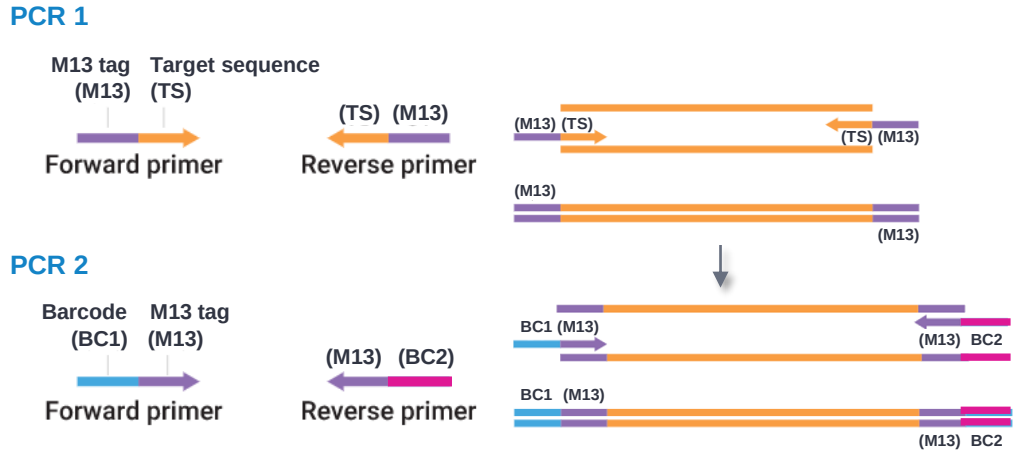
**PacBio barcoded M13 primer plate (102-135-500)**



**Procedure & checklist – Preparing multiplexed amplicon libraries using PacBio barcoded M13 primers and SMRTbell prep kit 3.0** (101-921-300)

- PacBio barcodes are added to amplicons through a 2-step PCR method using M13-tagged target-specific primers and M13-tagged 16-bp barcoded primers to produce asymmetrically barcoded samples
- Barcoded amplicons are pooled into a single tube for SMRTbell library construction
- Recommended for larger projects (up to 384 samples)

### 2-Step PCR



# Multiplexed amplicon sample preparation & sequencing workflow overview

Workflow summary for constructing multiplexed SMRTbell libraries suitable for sequencing on the Sequel and Sequel II and IIe systems for targeted amplicon sequencing applications

## PCR amplification

- If using barcoded primers for barcoding, follow *Procedure & checklist – Preparing multiplexed amplicon libraries using SMRTbell Prep Kit 3.0* (102-359-000)
- If using PacBio barcoded M13 primer plate (102-135-500), follow *Procedure & checklist – Preparing multiplexed amplicon libraries using PacBio barcoded M13 primers and SMRTbell prep kit 3.0* (101-921-300)
- Pool barcoded amplicons into a single tube and proceed with SMRTbell library construction

## SMRTbell library construction

- If using SMRTbell barcoded adapter plate 3.0 (102-009-200) to barcode your amplicons, follow *Procedure & checklist – Preparing multiplexed amplicon libraries using SMRTbell Prep Kit 3.0* (102-359-000) and pool adapter-barcoded samples into a single tube for SMRT sequencing

## Sequencing

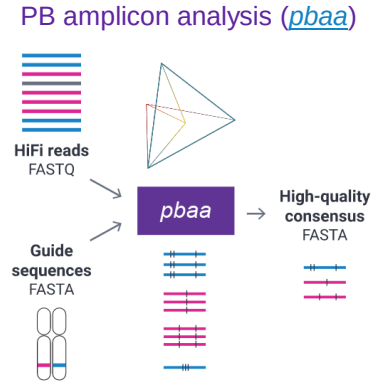
- Follow SMRT Link Sample Setup instructions for primer annealing, polymerase binding, complex cleanup and sample loading

HiFi Read

**PacBio HiFi reads achieve 99.9% accuracy**

## Data analysis

- Perform CCS analysis and de-multiplex barcodes on-instrument (Sequel IIe system only) or in SMRT Link
- Analyze demultiplexed HiFi data using SMRT Link or other command line tools

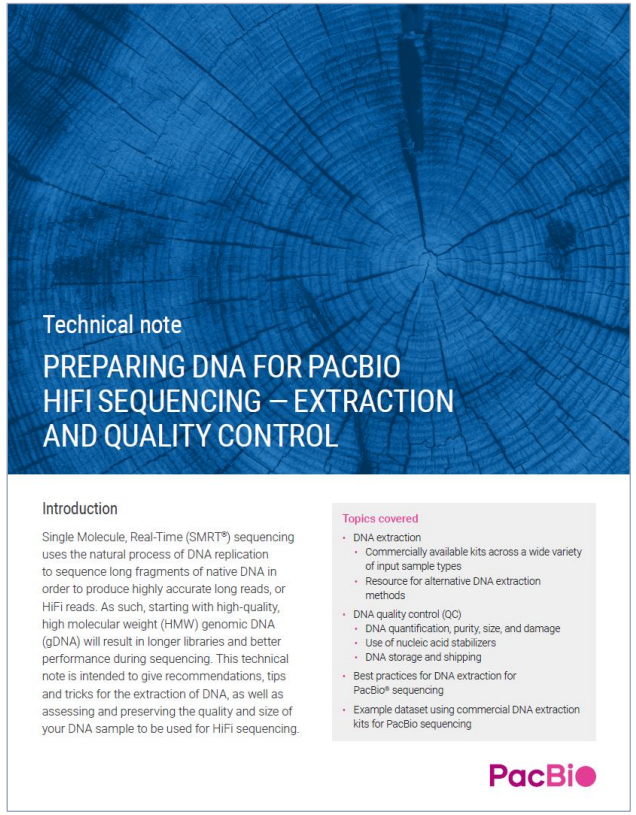


# Amplicon sample QC requirements

# Best practices for generating high-quality PCR amplicons

- Clean, target-specific PCR products are extremely important for obtaining high-quality sequence data.
- Non-specific products can represent a substantial percentage of the sequencing reads if they are not removed.
- To minimize their presence, consider the recommendations described in the following sections for generating high-quality amplicons suitable for SMRTbell library preparation and sequencing.

1. Begin with high-quality nucleic acids and work in a clean environment.
    - For targeted sequencing of genomic DNA (gDNA) samples, starting with high-quality DNA will result in better performance during sequencing.
        - PacBio Technical Note: DNA Prep (102-193-651) provides recommendations, tips and tricks for genomic DNA extraction as well as assessing and preserving the quality and size of your DNA sample to be used for PacBio HiFi sequencing
        - PacBio Technical Note: Sample preparation for PacBio HiFi sequencing from human whole blood (102-326-500) outlines best practices for extracting human whole blood samples using the Circulomics Nanobind CBB Big DNA Kit (NB-900-001-01)
    - If extracted nucleic acids must be stored, freeze at high concentrations in appropriately-buffered solutions.
    - To minimize possible contamination and degradation caused by multiple freeze/thaw cycles, sub-aliquot DNA into smaller volumes for storage.
    - Set up PCR reactions in an environment free from sources of non-specific primer and template contaminants, ideally a laminar flow hood, using dedicated pre-PCR pipettor, tips and reagents.




**PacBio Technical Note: Preparing DNA for PacBio HiFi sequencing - Extraction and quality control (102-193-651)**

# Best practices for generating high-quality PCR amplicons (cont.)

2. Use PCR reagents and conditions for generating target-specific, full-length amplicons.
    - Use the highest-fidelity polymerase compatible with your PCR amplification system.
    - Use desalted or HPLC-purified oligo primers; damaged bases at the ends of the amplicons cannot be repaired by DNA Damage Repair enzymes.
    - Optimize PCR conditions to minimize total time spent at high (>65°C) temperatures, particularly during denaturation.
    - PCR extension time should be long enough to ensure complete extension, taking into consideration the polymerase used and target amplicon size.
        - For mixed samples with similar targets, it is important to complete extension at every step to avoid generating chimeric products in subsequent steps.
        - As a general guideline, use extension times of one minute per 1000 base pairs (e.g., 3 minutes for a 3 kb product).
3. Use the lowest number of cycles required for obtaining adequate yields (ng) of PCR products to proceed with SMRTbell library construction. Avoid over-amplification.
4. If non-specific products are present, optimize PCR conditions or perform AMPure PB bead-based size selection to enrich for PCR amplicons with the desired target size
5. Note: Using gel-extracted amplicon products may result in lower sequencing performance due to the damage inherently caused by intercalating dyes such as ethidium bromide and exposure to UV radiation.
    - Sequencing amplicons stained with SYBR dyes from ThermoFisher Scientific is untested, and thus, also **not** recommended.
    - If working with a gel product that has been stained with a dye, it is recommended to bring it through **additional rounds of amplification** to remove damage and dyes prior to library prep and sequencing.

# Evaluation of PCR amplicon DNA concentration

- For accurate quantification of PCR samples to be used in multiplexed amplicon library preparation workflows, PacBio recommends using the Qubit fluorometer and Qubit dsDNA High Sensitivity (HS) assay reagents (Thermo Fisher Scientific)
  - Qubit dsDNA HS assay quantitation range: 0.2 – 100 ng
  - **Note:** When measuring very low DNA concentrations of amplicon samples (especially <1 kb library insert size), it may be helpful to **increase the sample aliquot volume above 1 µL** (up to 20 µL) in order to ensure sufficient assay sensitivity

**Qubit dsDNA HS Assay Kit**


**Qubit 4 Fluorometer**

# Evaluation of PCR amplicon DNA purity

- DNA purity can be determined by using a NanoDrop system [Thermo Fisher Scientific] or other spectrophotometer tool
- For ultrapure DNA, A260/280 ratio is typically between ~1.8 - 2.0 and A260/230 ratio is ≥2.0
- **Note:** High UV absorbance values are *not* always a guarantee of optimal sequencing performance because not all inhibitors absorb at the wavelengths of 230, 260, and 280 nm.
- Conversely, low UV absorbance values are *not* always a guarantee that non-optimal sequencing performance will be obtained for a sample

## A260/A280 Ratio

- A low A260/A280 ratio may be the result of:
  - Protein
  - Phenol
  - Other contaminants that absorb strongly at or near 280 nm
  - Sometimes it may be caused by a very low concentration of nucleic acid.
- High 260/280 ratios are not indicative of an issue

If A260/280 and A260/230 readings are out of the recommended ranges, perform one or more rounds of **AMPure PB bead purification** followed by re-assessment of the quantity and purity of the input DNA sample.**

## A260/A230 Ratio

- A low A260/A230 ratio may be the result of:
  - Protein*
  - Carbohydrate carryover (often a problem with plants)
  - Residual phenol from nucleic acid extraction
  - Residual guanidine (often used in column-based kits)
  - Glycogen used for precipitation
- A high A260/A230 ratio may be the result of:
  - Making a blank measurement on a dirty pedestal of a Nanodrop instrument
  - Using an inappropriate solution for the blank measurement


* See NEB Technical Note: A Practical Guide to Analyzing Nucleic Acid Concentration and Purity with Microvolume Spectrophotometers (2019)

** See PacBio Technical Note: Preparing DNA for PacBio HiFi sequencing - Extraction and quality control (102-193-651)

# Evaluation of amplicon DNA size distribution

- It is important to accurately assess the sizes of the amplicons that are being multiplexed before preparing SMRTbell libraries for sequencing
- For sizing QC of amplicons, visualize an aliquot of each PCR reaction using an Agilent Bioanalyzer system, Agilent TapeStation system, Agilent Fragment Analyzer system, Agilent Femto Pulse system or manual agarose gel electrophoresis with appropriate markers or ladders
- If off-target/non-specific products are present, optimize PCR conditions or perform one or more rounds of AMPure PB bead-based purification to enrich for PCR amplicons with the desired target size.
    - If the contaminating bands are quite close in size or larger than the desired amplicon, or for any contaminants >1.5 kb, a gel-based size selection method is recommended

- Pool amplicons ≤3 kb separately from those >3 kb in length for optimal loading and sequencing performance
- Pooling amplicons of significantly different sizes together will increase sequence coverage variability because of differences in molarity between those samples.
- Differences in the number of molecules in a sample will translate into differences in the number of molecules loaded and sequenced on a SMRT Cell.

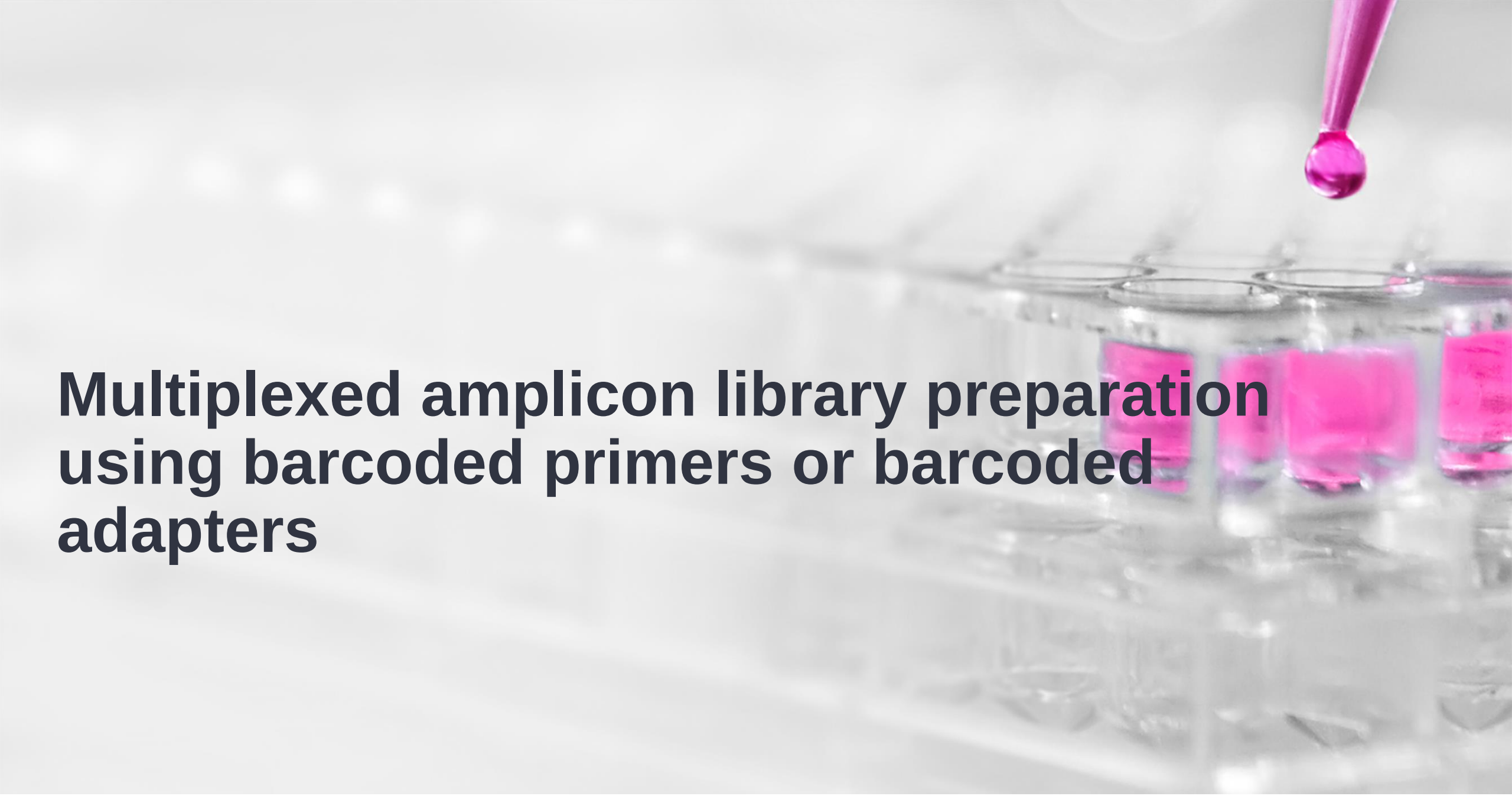
Bioanalyzer 2100 System (Agilent Technologies)

4200 TapeStation System (Agilent Technologies)

Fragment Analyzer System (Agilent Technologies)

Femto Pulse System (Agilent Technologies)

# Multiplexed amplicon library preparation using barcoded primers or barcoded adapters

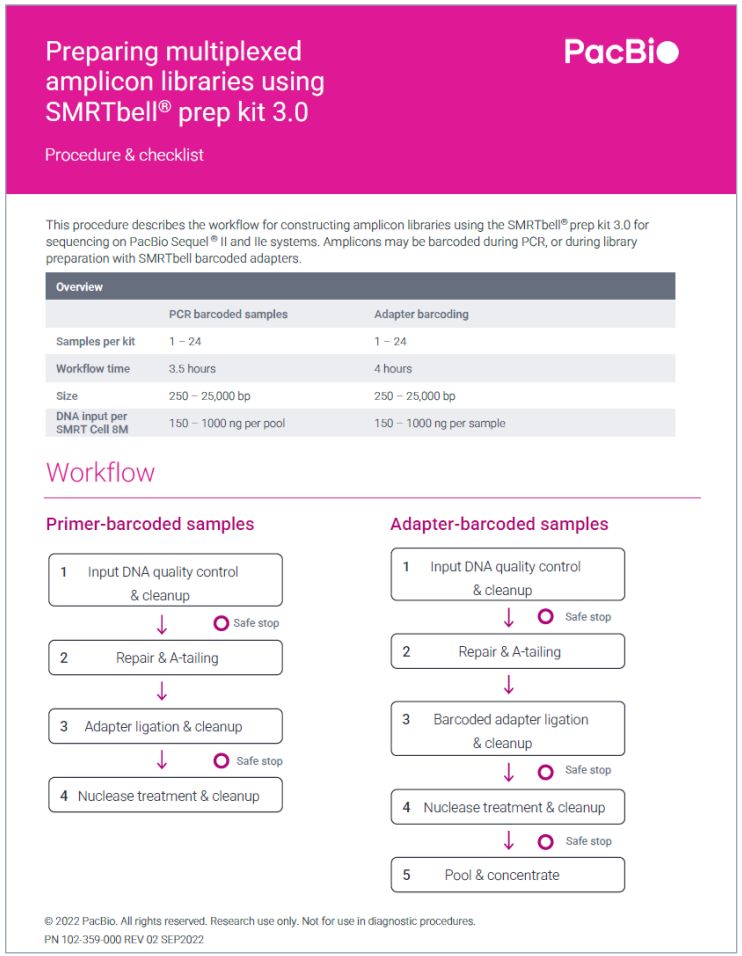


# Procedure & checklist – Preparing multiplexed amplicon libraries using SMRTbell prep kit 3.0 (102-359-000)

Procedure & checklist 102-359-000 describes the workflow for constructing amplicon libraries using the SMRTbell prep kit 3.0 for sequencing on PacBio Sequel II and IIe Systems for targeted sequencing applications. Amplicons may be barcoded during PCR using barcoded primers or during library construction with SMRTbell barcoded adapters, or by both methods to create nested barcode combinations.

## Procedure & checklist contents

1. General best practices for input amplicon DNA QC.
2. Multiplexing best practices guidance for pooling amplicon samples for SMRTbell library construction and SMRT sequencing
3. Enzymatic workflow steps for preparation of multiplexed SMRTbell libraries using the SMRTbell prep kit 3.0 (102-182-700) and (if barcoding amplicons during library construction) the SMRTbell barcoded adapter plate 3.0 (102-009-200).



Preparing multiplexed amplicon libraries using SMRTbell® prep kit 3.0

PacBio

Procedure & checklist

This procedure describes the workflow for constructing amplicon libraries using the SMRTbell® prep kit 3.0 for sequencing on PacBio Sequel® II and IIe systems. Amplicons may be barcoded during PCR, or during library preparation with SMRTbell barcoded adapters.

| Overview | | |
|---|---|---|
| | PCR barcoded samples | Adapter barcoding |
| Samples per kit | 1 – 24 | 1 – 24 |
| Workflow time | 3.5 hours | 4 hours |
| Size | 250 – 25,000 bp | 250 – 25,000 bp |
| DNA input per SMRT Cell 8M | 150 – 1000 ng per pool | 150 – 1000 ng per sample |

Workflow

Primer-barcoded samples

1 Input DNA quality control & cleanup
Safe stop
2 Repair & A-tailing
3 Adapter ligation & cleanup
Safe stop
4 Nuclease treatment & cleanup

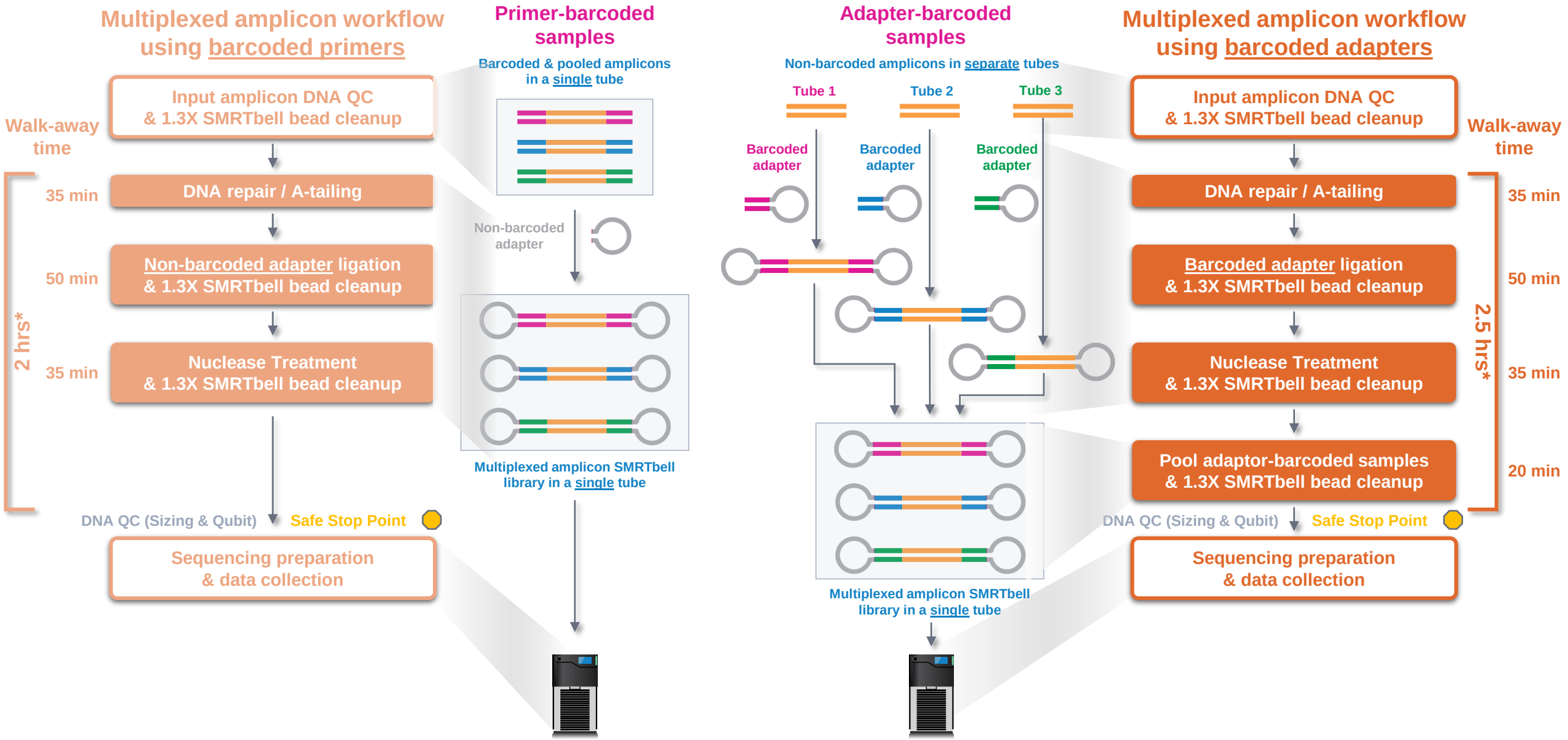
Adapter-barcoded samples

1 Input DNA quality control & cleanup
Safe stop
2 Repair & A-tailing
3 Barcoded adapter ligation & cleanup
Safe stop
4 Nuclease treatment & cleanup
Safe stop
5 Pool & concentrate


PN 102-359-000 REV 02 SEP2022

PacBio Documentation (102-359-000)

# Amplicon SMRTbell library construction workflow overview using barcoded primers or barcoded adapters




* **Excludes** hands-on time and additional time required to perform DNA Sizing QC and DNA concentration QC

# SMRTbell barcoded adapter plate 3.0

## For Sequel II and IIe systems, SMRTbell barcoded adapter plate 3.0 (102-009-200) is available for multiplexing amplicon samples

- Use barcoded adapters from SMRTbell barcoded adapter plate 3.0 for barcoding amplicon samples at Step 3 (“Adapter ligation & cleanup”) in the procedure
  - Pooling of adapter-barcoded libraries is described in Step 5 of the protocol
- SMRTbell barcoded adapter plate 3.0 contains 96 barcoded adapters to support multiplexed SMRTbell library construction for up to 96 samples using SMRTbell prep kit 3.0
  - Each barcoded adapter contains a 5 bp padding sequence for more uniform ligation performance across different barcode sequences
  - Each well on the plate contains a barcoded adapter with a unique 10-base pair PacBio barcode sequence
  - Each barcoded adapter is present in only one well and supports a single reaction
- SMRT Link comes pre-installed with the following barcode set FASTA file containing SMRTbell barcoded adapter plate 3.0 barcode sequences*:
  `SMRTbell Barcoded Adapter Plate 3.0 (bc2001-bc2096)`

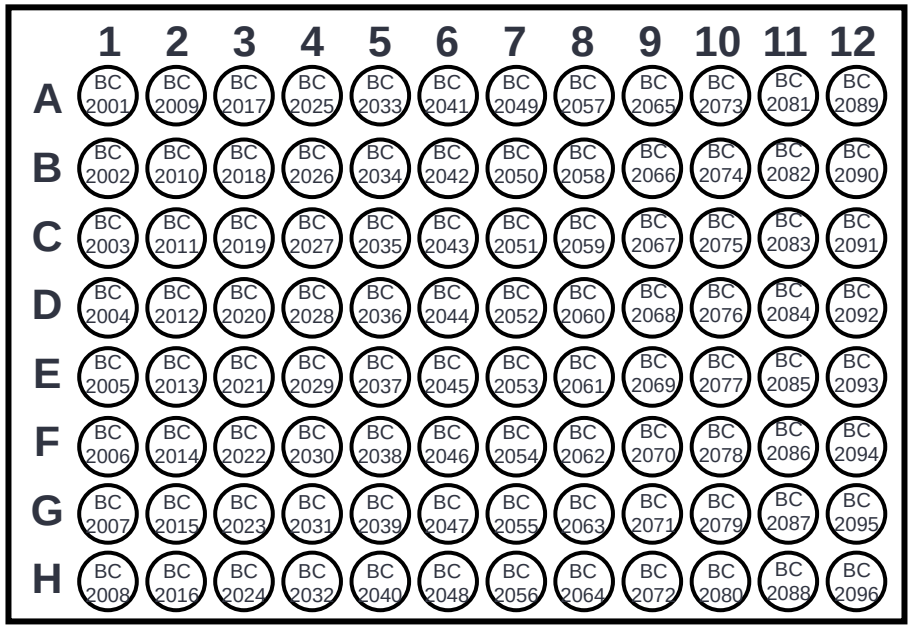



**Figure illustration of mapping between a specific well location and a unique PacBio barcode sequence on a 96-well plate in the SMRTbell Barcoded Adapter Plate (102-009-200)**

Reagent kit quantities support a **single use** of each of the 96 barcoded adapters in the plate for SMRTbell library preparations.

Plate Layout (Excel) [ Link ]

Barcode Sequences (FASTA) [ Link ]

Product insert: SMRTbell barcoded adapter plate 3.0 (96 barcodes, 96 samples) [ Link ]

* If your SMRT Link version does not have the required barcode set FASTA file for a specific analysis application, please visit PacBio’s Multiplexing website to download the appropriate FASTA file or contact PacBio Technical Support.

# General best practices for multiplexed amplicon library preparation using barcoded primers or barcoded adapters

## Amplicon DNA input requirements for SMRTbell library construction using barcoded PCR primers

The total amount of DNA required for constructing a SMRTbell library is dependent on the mean size of the amplicons being sequenced as shown in the table below

### For samples multiplexed with barcoded PCR primers:

- Samples can be pooled prior to library preparation and the total input amplicon DNA amount will equal the total combined mass of the multiplexed amplicon pool
  - The per-sample input will equal the total DNA input divided by the number of multiplexed samples
  - E.g., for a 96-plex of >7 kb amplicon samples, the minimum required per-sample amount would be ~3.2 ng (≈300 ng / 96)
- Use no less than 150 ng of total input per SMRT Cell 8M to ensure sufficient library yields for optimal SMRT cell loading when working with targets less than a mean size of 5 kb
  - Larger amplicons will require more input material to achieve desired molarity for SMRT cell loading
  - Refer to the table for the recommended minimum total input amounts per SMRT Cell 8M that are required for SMRTbell library construction

| Mean amplicon size | Minimum total pooled input DNA amount per SMRT Cell 8M* |
|---|---|
| **<5 kb** | 150 ng |
| **5 – 7 kb** | 200 ng |
| **>7 kb** | 300 ng |

PacBio

* Lower total input amounts are possible but may result in lower sequencing yield if there is not enough SMRTbell library available for optimal SMRT cell loading.

# General best practices for multiplexed amplicon library preparation using barcoded primers or barcoded adapters

## Amplicon DNA input requirements for SMRTbell library construction using barcoded adapters

The total amount of DNA required for constructing a SMRTbell library is dependent on the mean size of the amplicons being sequenced as shown in the table below

### For samples multiplexed with barcoded adapters:

- Use a minimum of 150 ng of DNA input per sample for samples with mean size less than 5 kb (use ≥200 ng per sample for amplicons between 5 and 7 kb, and use ≥300 ng per sample for amplicons greater than 7 kb)
  - This is to ensure sufficient recovery of each sample at the end of library prep for equal mass or equal molar pooling
- Using lower per-sample amounts, though possible, may result in low library yields and lead to uneven pooling and sequence coverage
  - For applications that require lower input amounts, consider using **barcoded primers** so samples can be pooled prior to library prep

| Mean amplicon size | Minimum per-sample input DNA amount per SMRT Cell 8M* |
|---|---|
| **<5 kb** | 150 ng |
| **5 – 7 kb** | 200 ng |
| **>7 kb** | 300 ng |



* Using lower per-sample amounts, though possible, may result in low library yields and lead to uneven pooling and sequence coverage.

# General best practices for multiplexed amplicon library preparation using barcoded primers or barcoded adapters (cont.)

## Multiplexing and pooling best practices

- When working with a large number of reactions, we recommend using a multichannel pipette to transfer small aliquots of master mixes to a 96-well or 384-well plate
  - Prepare master mixes according to the instructions in the procedure.
  - Transfer aliquots of the master mix into an 8-tube strip using a single channel pipette (1/8th master mix volume to each of the eight well of the strip tube). Each tube can accommodate up to 200 µL of liquid.
  - Using an 8-channel pipette, transfer the required reaction volume of the master mix from the 8-tube strip into the appropriate sample wells of a 96-well or 384-well plate.
  - Repeat until all required reaction wells in the sample plate are filled.
- Use the SMRTbell barcoded adapter plate 3.0 when barcoding samples using a barcoded SMRTbell adapter. Quick spin the plate to collect liquid at bottom of the well prior to use.
- Pool amplicons of similar size for optimal sequencing performance.
  - Pool amplicons ≤3 kb separately from amplicons >3 kb for optimal sequencing yields across all samples.
- When amplicons are similar in size, pool an equal mass for each sample. Some experiments may require equal molar pooling if the mean size differs between samples and similar coverage levels are required.
- Pooling amplicons of different sizes together will increase sequence coverage variability because of differences in molarity between those samples.
  - Differences in the number of molecules in a sample will translate to the differences in the number of molecules loaded and sequenced on the SMRT Cell 8M

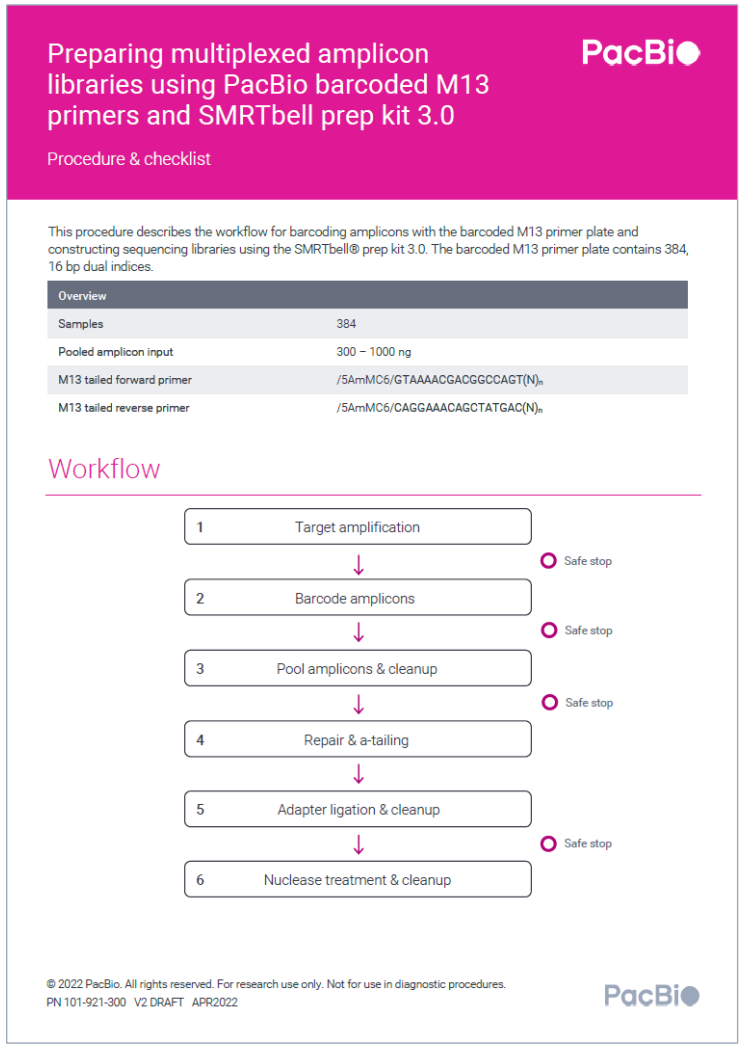
# Multiplexed amplicon library preparation using barcoded M13 primers

# Procedure & checklist – Preparing multiplexed amplicon libraries using PacBio barcoded M13 primers and SMRTbell prep kit 3.0 (101-921-300)

Procedure & checklist 101-921-300 describes a method for constructing SMRTbell libraries using the SMRTbell prep kit 3.0 (SPK 3.0) and the barcoded M13 primer plate that are suitable for generating HiFi reads on the PacBio Sequel II and IIe systems for targeted amplicon sequencing applications

## Procedure & checklist contents

1. General best practices for PCR reagent handling and PCR optimization.
2. 2-step PCR workflow for 1) target amplification and 2) barcoding amplicons with the barcoded M13 primer plate (102-135-500).
3. Multiplexing best practices guidance for pooling barcoded amplicons for SMRTbell library construction.
4. Enzymatic workflow steps for preparation of SMRTbell libraries from barcoded amplicon products using SMRTbell prep kit 3.0 (102-182-700).



Preparing multiplexed amplicon libraries using PacBio barcoded M13 primers and SMRTbell prep kit 3.0

PacBio

Procedure & checklist

This procedure describes the workflow for barcoding amplicons with the barcoded M13 primer plate and constructing sequencing libraries using the SMRTbell® prep kit 3.0. The barcoded M13 primer plate contains 384, 16 bp dual indices.

| Overview | |
|---|---|
| Samples | 384 |
| Pooled amplicon input | 300 – 1000 ng |
| M13 tailed forward primer | /5AmMC6/GTAAAACGACGGCCAGT(N)$_n$ |
| M13 tailed reverse primer | /5AmMC6/CAGGAAACAGCTATGAC(N)$_n$ |

Workflow




PN 101-921-300 V2 DRAFT APR2022

PacBio

PacBio Documentation (101-921-300)

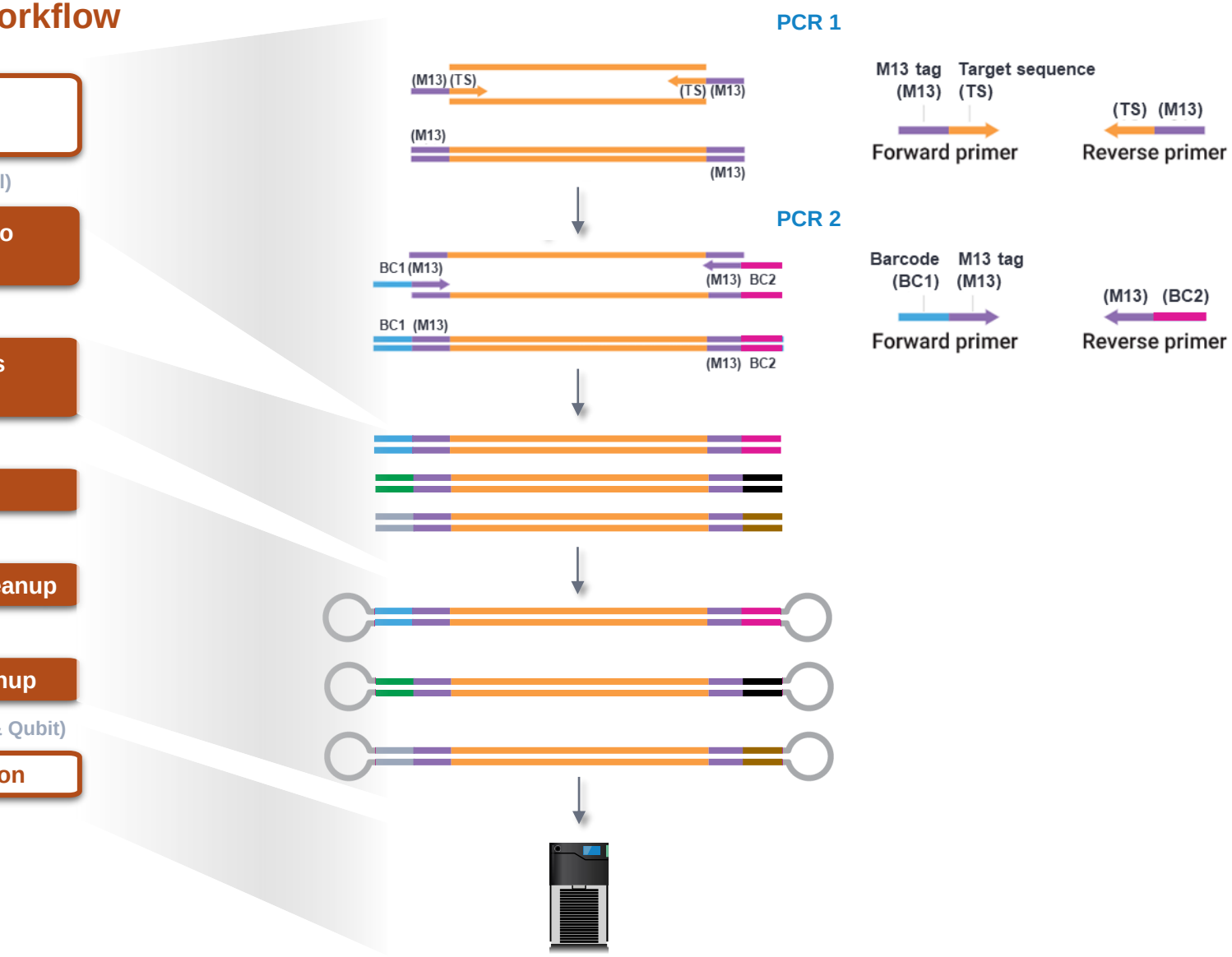
# Amplicon SMRTbell library construction workflow overview using barcoded M13 primers

## SPK 3.0 BC_M13P Amplicon Workflow

(101-921-300)

Walk-away time ~7 hours*

| | |
|---|---|
| 130 min | PCR 1: Target amplification with M13-tailed gene-specific primers |
| | DNA QC (Optional) |
| 130 min | PCR 2: Barcode amplicons with PacBio barcoded M13 primers |
| | DNA QC (Sizing) |
| 20 min | Equal-volume pool barcoded samples & 1.3X SMRTbell bead cleanup |
| | Safe Stop Point — DNA QC (Qubit) |
| 35 min | DNA repair / A-tailing |
| 50 min | Adapter ligation & 1.3X SMRTbell bead cleanup |
| 35 min | Nuclease Tx & 1.3X SMRTbell bead cleanup |
| | Safe Stop Point — DNA QC (Sizing & Qubit) |
| | Sequencing preparation & data collection |




* **Excludes** hands-on time and additional time required to perform DNA Sizing QC and DNA concentration QC

# Barcoded M13 Primer Plate

## Asymmetric barcode plate map for Barcoded M13 primer plate (102-135-500)

- Ready-to-use premixed primer plate containing 384 barcoded M13 primer pairs for asymmetric (dual index) barcoding of multiplexed SMRTbell libraries
  - Plate includes 40 different oligos (16 M13 forward primers + 24 M13 reverse primers)
- Single-use per well with pierceable foil (can reseal between sample batches)
  - Fill volume in each well = 12 µl (at 10 µM primer concentration)
- Plate Layout (Excel): Link
- Barcode Sequences (FASTA): Link

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 1002 1050 | 1002 1051 | 1002 1052 | 1002 1053 | 1002 1054 | 1002 1055 | 1002 1056 | 1002 1057 | 1002 1058 | 1002 1059 | 1002 1060 | 1002 1061 | 1002 1062 | 1002 1063 | 1002 1064 | 1002 1065 | 1002 1066 | 1002 1067 | 1002 1068 | 1002 1069 | 1002 1070 | 1002 1071 | 1002 1072 | 1002 1073 |
| B | 1003 1050 | 1003 1051 | 1003 1052 | 1003 1053 | 1003 1054 | 1003 1055 | 1003 1056 | 1003 1057 | 1003 1058 | 1003 1059 | 1003 1060 | 1003 1061 | 1003 1062 | 1003 1063 | 1003 1064 | 1003 1065 | 1003 1066 | 1003 1067 | 1003 1068 | 1003 1069 | 1003 1070 | 1003 1071 | 1003 1072 | 1003 1073 |
| C | 1004 1050 | 1004 1051 | 1004 1052 | 1004 1053 | 1004 1054 | 1004 1055 | 1004 1056 | 1004 1057 | 1004 1058 | 1004 1059 | 1004 1060 | 1004 1061 | 1004 1062 | 1004 1063 | 1004 1064 | 1004 1065 | 1004 1066 | 1004 1067 | 1004 1068 | 1004 1069 | 1004 1070 | 1004 1071 | 1004 1072 | 1004 1073 |
| D | 1005 1050 | 1005 1051 | 1005 1052 | 1005 1053 | 1005 1054 | 1005 1055 | 1005 1056 | 1005 1057 | 1005 1058 | 1005 1059 | 1005 1060 | 1005 1061 | 1005 1062 | 1005 1063 | 1005 1064 | 1005 1065 | 1005 1066 | 1005 1067 | 1005 1068 | 1005 1069 | 1005 1070 | 1005 1071 | 1005 1072 | 1005 1073 |
| E | 1006 1050 | 1006 1051 | 1006 1052 | 1006 1053 | 1006 1054 | 1006 1055 | 1006 1056 | 1006 1057 | 1006 1058 | 1006 1059 | 1006 1060 | 1006 1061 | 1006 1062 | 1006 1063 | 1006 1064 | 1006 1065 | 1006 1066 | 1006 1067 | 1006 1068 | 1006 1069 | 1006 1070 | 1006 1071 | 1006 1072 | 1006 1073 |
| F | 1007 1050 | 1007 1051 | 1007 1052 | 1007 1053 | 1007 1054 | 1007 1055 | 1007 1056 | 1007 1057 | 1007 1058 | 1007 1059 | 1007 1060 | 1007 1061 | 1007 1062 | 1007 1063 | 1007 1064 | 1007 1065 | 1007 1066 | 1007 1067 | 1007 1068 | 1007 1069 | 1007 1070 | 1007 1071 | 1007 1072 | 1007 1073 |
| G | 1008 1050 | 1008 1051 | 1008 1052 | 1008 1053 | 1008 1054 | 1008 1055 | 1008 1056 | 1008 1057 | 1008 1058 | 1008 1059 | 1008 1060 | 1008 1061 | 1008 1062 | 1008 1063 | 1008 1064 | 1008 1065 | 1008 1066 | 1008 1067 | 1008 1068 | 1008 1069 | 1008 1070 | 1008 1071 | 1008 1072 | 1008 1073 |
| H | 1009 1050 | 1009 1051 | 1009 1052 | 1009 1053 | 1009 1054 | 1009 1055 | 1009 1056 | 1009 1057 | 1009 1058 | 1009 1059 | 1009 1060 | 1009 1061 | 1009 1062 | 1009 1063 | 1009 1064 | 1009 1065 | 1009 1066 | 1009 1067 | 1009 1068 | 1009 1069 | 1009 1070 | 1009 1071 | 1009 1072 | 1009 1073 |
| I | 1010 1050 | 1010 1051 | 1010 1052 | 1010 1053 | 1010 1054 | 1010 1055 | 1010 1056 | 1010 1057 | 1010 1058 | 1010 1059 | 1010 1060 | 1010 1061 | 1010 1062 | 1010 1063 | 1010 1064 | 1010 1065 | 1010 1066 | 1010 1067 | 1010 1068 | 1010 1069 | 1010 1070 | 1010 1071 | 1010 1072 | 1010 1073 |
| J | 1011 1050 | 1011 1051 | 1011 1052 | 1011 1053 | 1011 1054 | 1011 1055 | 1011 1056 | 1011 1057 | 1011 1058 | 1011 1059 | 1011 1060 | 1011 1061 | 1011 1062 | 1011 1063 | 1011 1064 | 1011 1065 | 1011 1066 | 1011 1067 | 1011 1068 | 1011 1069 | 1011 1070 | 1011 1071 | 1011 1072 | 1011 1073 |
| K | 1012 1050 | 1012 1051 | 1012 1052 | 1012 1053 | 1012 1054 | 1012 1055 | 1012 1056 | 1012 1057 | 1012 1058 | 1012 1059 | 1012 1060 | 1012 1061 | 1012 1062 | 1012 1063 | 1012 1064 | 1012 1065 | 1012 1066 | 1012 1067 | 1012 1068 | 1012 1069 | 1012 1070 | 1012 1071 | 1012 1072 | 1012 1073 |
| L | 1013 1050 | 1013 1051 | 1013 1052 | 1013 1053 | 1013 1054 | 1013 1055 | 1013 1056 | 1013 1057 | 1013 1058 | 1013 1059 | 1013 1060 | 1013 1061 | 1013 1062 | 1013 1063 | 1013 1064 | 1013 1065 | 1013 1066 | 1013 1067 | 1013 1068 | 1013 1069 | 1013 1070 | 1013 1071 | 1013 1072 | 1013 1073 |
| M | 1014 1050 | 1014 1051 | 1014 1052 | 1014 1053 | 1014 1054 | 1014 1055 | 1014 1056 | 1014 1057 | 1014 1058 | 1014 1059 | 1014 1060 | 1014 1061 | 1014 1062 | 1014 1063 | 1014 1064 | 1014 1065 | 1014 1066 | 1014 1067 | 1014 1068 | 1014 1069 | 1014 1070 | 1014 1071 | 1014 1072 | 1014 1073 |
| N | 1015 1050 | 1015 1051 | 1015 1052 | 1015 1053 | 1015 1054 | 1015 1055 | 1015 1056 | 1015 1057 | 1015 1058 | 1015 1059 | 1015 1060 | 1015 1061 | 1015 1062 | 1015 1063 | 1015 1064 | 1015 1065 | 1015 1066 | 1015 1067 | 1015 1068 | 1015 1069 | 1015 1070 | 1015 1071 | 1015 1072 | 1015 1073 |
| O | 1016 1050 | 1016 1051 | 1016 1052 | 1016 1053 | 1016 1054 | 1016 1055 | 1016 1056 | 1016 1057 | 1016 1058 | 1016 1059 | 1016 1060 | 1016 1061 | 1016 1062 | 1016 1063 | 1016 1064 | 1016 1065 | 1016 1066 | 1016 1067 | 1016 1068 | 1016 1069 | 1016 1070 | 1016 1071 | 1016 1072 | 1016 1073 |
| P | 1017 1050 | 1017 1051 | 1017 1052 | 1017 1053 | 1017 1054 | 1017 1055 | 1017 1056 | 1017 1057 | 1017 1058 | 1017 1059 | 1017 1060 | 1017 1061 | 1017 1062 | 1017 1063 | 1017 1064 | 1017 1065 | 1017 1066 | 1017 1067 | 1017 1068 | 1017 1069 | 1017 1070 | 1017 1071 | 1017 1072 | 1017 1073 |

FORWARD
Reverse

# General best practices for multiplexed amplicon library preparation using barcoded M13 primers

## PCR best practices

- Add a 5’ block (\5AmMC6\) and M13 sequence to all first-round, target-specific primers. The 5’ block prevents unbarcoded amplicons from ligating to the SMRTbell adapters during library prep
- Follow the manufacturer’s instructions and any necessary adjustments to annealing temperature, $MgCl_2$ concentration, and GC-rich targets to optimize PCR
- Keep all KAPA HiFi HotStart reagents and reactions on ice until PCR; the high proofreading activity of the enzyme will rapidly degrade primers at room temperature. This is generally true for all high-fidelity polymerases
- Use high-quality DNA and work in a PCR-clean environment to avoid contamination
- Use a non-template control (NTC) to check for contamination
- Optimize PCR parameters to enable equal volume pooling and prevent off-target amplification and primer-dimers. Off-target products and high levels of primer dimers may reduce sequencing yields and performance
- Use the fewest number of PCR cycles required for obtaining adequate yields (ng)
- Avoid using gel-extraction and intercalating dyes such as ethidium bromide on the 2$^{nd}$ round (barcoded) amplicons because this causes DNA damage which will impact sequencing yields

# General best practices for multiplexed amplicon library preparation using barcoded M13 primers (cont.)

## Amplicon DNA input requirements for SMRTbell library construction

The total amount of DNA required for constructing a SMRTbell library is dependent on the mean size of the amplicons being sequenced as shown in the table below

### For samples multiplexed with barcoded M13 primers:

- The total input DNA amount per SMRT Cell 8M required for SMRTbell library construction should equal the total combined mass of the multiplexed amplicon pool
    - E.g., for >10 kb amplicon samples, the total required pooled input DNA amount would be 1000 ng
- Use a total pooled amplicon amount of 300 – 1000 ng to ensure optimal loading and sequencing yields
- Larger amplicons require higher input amounts relative to smaller amplicons to achieve the required molarity for SMRT cell loading

| Amplicon size | Total pooled amplicon DNA input per SMRT Cell 8M |
|---|---|
| **<1.5 kb** | 300 ng |
| **1.5 – 3 kb** | 300 ng |
| **3 – 10 kb** | 500 ng |
| **≥10 kb** | 1000 ng |

# General best practices for multiplexed amplicon library preparation using barcoded M13 primers (cont.)

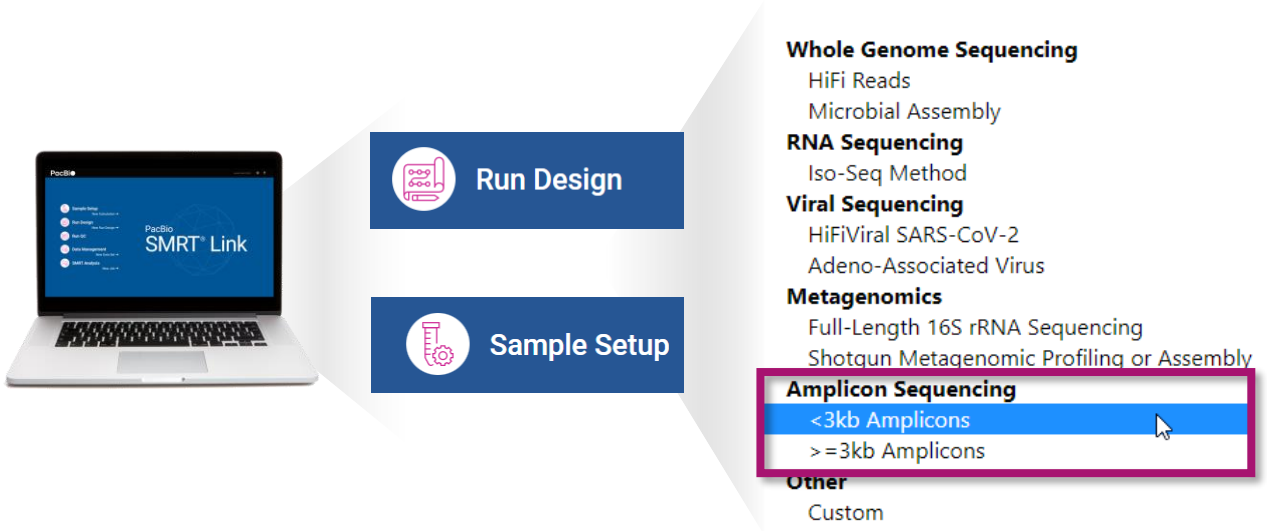
## Multiplexing and pooling best practices

- When working with a large number of reactions, we recommend using a multichannel pipette to transfer small aliquots of master mixes to a 96-well or 384-well plate
  - Prepare master mixes according to the instructions in the procedure.
  - Transfer aliquots of the master mix into an 8-tube strip using a single channel pipette (1/8th master mix volume to each of the eight well of the strip tube). Each tube can accommodate up to 200 µL of liquid.
  - Using an 8-channel pipette, transfer the required reaction volume of the master mix from the 8-tube strip into the appropriate sample wells of a 96-well or 384-well plate.
  - Repeat until all required reaction wells in the sample plate are filled.
- Pool amplicons of similar size for optimal sequencing performance.
  - Pool amplicons ≤3 kb separately from those >3 kb in length for optimal sequencing yields across all samples
- Normalizing DNA input into PCR and optimizing PCR will all improve sequence-coverage balance across samples when pooling amplicons in an equal volume fashion
- Pooling amplicons of different sizes together will increase sequence coverage variability because of differences in molarity between those samples.
  - Differences in the number of molecules in a sample will translate to the differences in the number of molecules loaded and sequenced on the SMRT Cell 8M

# Multiplexed amplicon library sequencing preparation workflow overview

# Sample Setup & Run Design recommendations for amplicon libraries

## In SMRT Link Sample Setup & Run Design, select ‘Amplicon Sequencing’ and choose ‘<3 kb Amplicons’ or ‘≥3 kb Amplicons’ for application type




- After specifying your application type, **SMRT Link** auto-fills selected Sample Setup and Run Design parameter fields with default recommended values*

| Amplicon library type | Recommended binding kit |
|---|---|
| Amplicons <3 kb | Sequel II binding kit 3.1 |
| Amplicons ≥3 kb | Sequel II binding kit 3.2 |


**Sequel II binding kit 3.1 & cleanup beads (102-333-400) is recommended for preparing <3 kb amplicon samples for sequencing.**

**Sequel II binding kit 3.1 & cleanup beads** (102-333-400) includes:

- Sequencing primer 3.1
- Sequel II polymerase 2.1
- DNA internal control 3.1 (defined 2 kb template bound to Polymerase 2.1)
- SMRTbell cleanup beads for complex cleanup


**Sequel II binding kit 3.2 & cleanup beads (102-333-300) is recommended for preparing ≥3 kb amplicon samples for sequencing.**

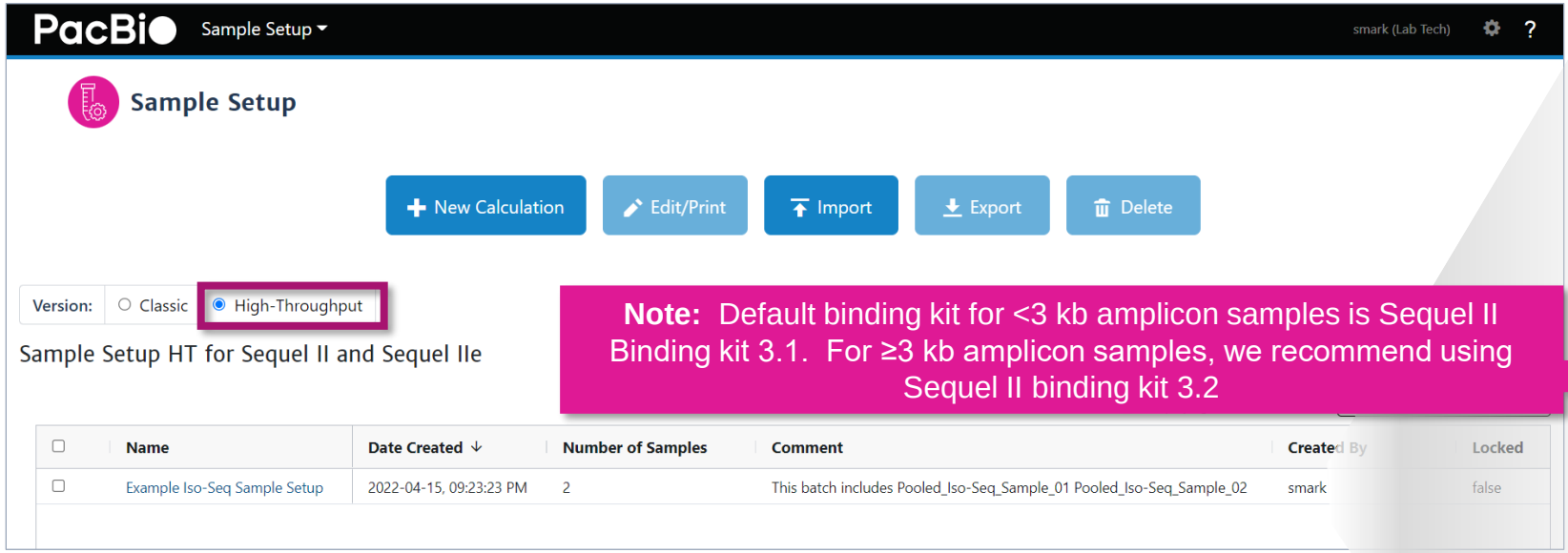
**Sequel II binding kit 3.2 & cleanup beads** (102-333-300) includes:

- Sequencing primer 3.2
- Sequel II polymerase 2.2
- DNA internal control 3.2 (defined 11 kb template bound to Polymerase 2.2)
- SMRTbell cleanup beads for complex cleanup

* See **Quick reference card – loading and pre-extension recommendations for Sequel II and IIe systems** (101-769-100) for any updates to sample setup and run design recommendations for specific applications.

# Amplicon library Sample Setup guidance

Use SMRT Link Sample Setup High-Throughput (HT) mode and follow instructions to perform ABC (anneal primer / bind polymerase / clean up complex) using recommended settings for amplicon samples




**Note:** Default binding kit for <3 kb amplicon samples is Sequel II Binding kit 3.1. For ≥3 kb amplicon samples, we recommend using Sequel II binding kit 3.2

- Sample Setup High-Throughput mode provides a simplified, streamlined workflow to efficiently process either one sample or multiple samples with similar library properties (such as mean insert size and DNA concentration) in parallel
- You can also export the calculated values to a CSV file for laboratory automation

| | Sample Group |
|---|---|
| Actions | Copy / Remove / Lock / Automate |
| Name | Example Amplicon Sample Setup |
| Comment | This batch includes: Pooled_Amplicon_Sample_01 Pooled_Amplicon_Sample_01 |
| Application | <3kb Amplicons |
| Binding Kit | Sequel® II Binding Kit 3.1 |
| Number of Samples | 2 samples |
| SMRT Cells per Sample | 1 cells |
| Available Volume per Sample | 15 uL |
| Insert Size | 2800 bp |
| Sample Concentration | 10 ng/uL |
| Cleanup Anticipated Yield | 60 % |
| Recommended Concentration on Plate | 40-150 pM |
| Specify Concentration on Plate | 125 pM |
| Minimum Pipetting Volume | 1 uL |
| Warnings | |

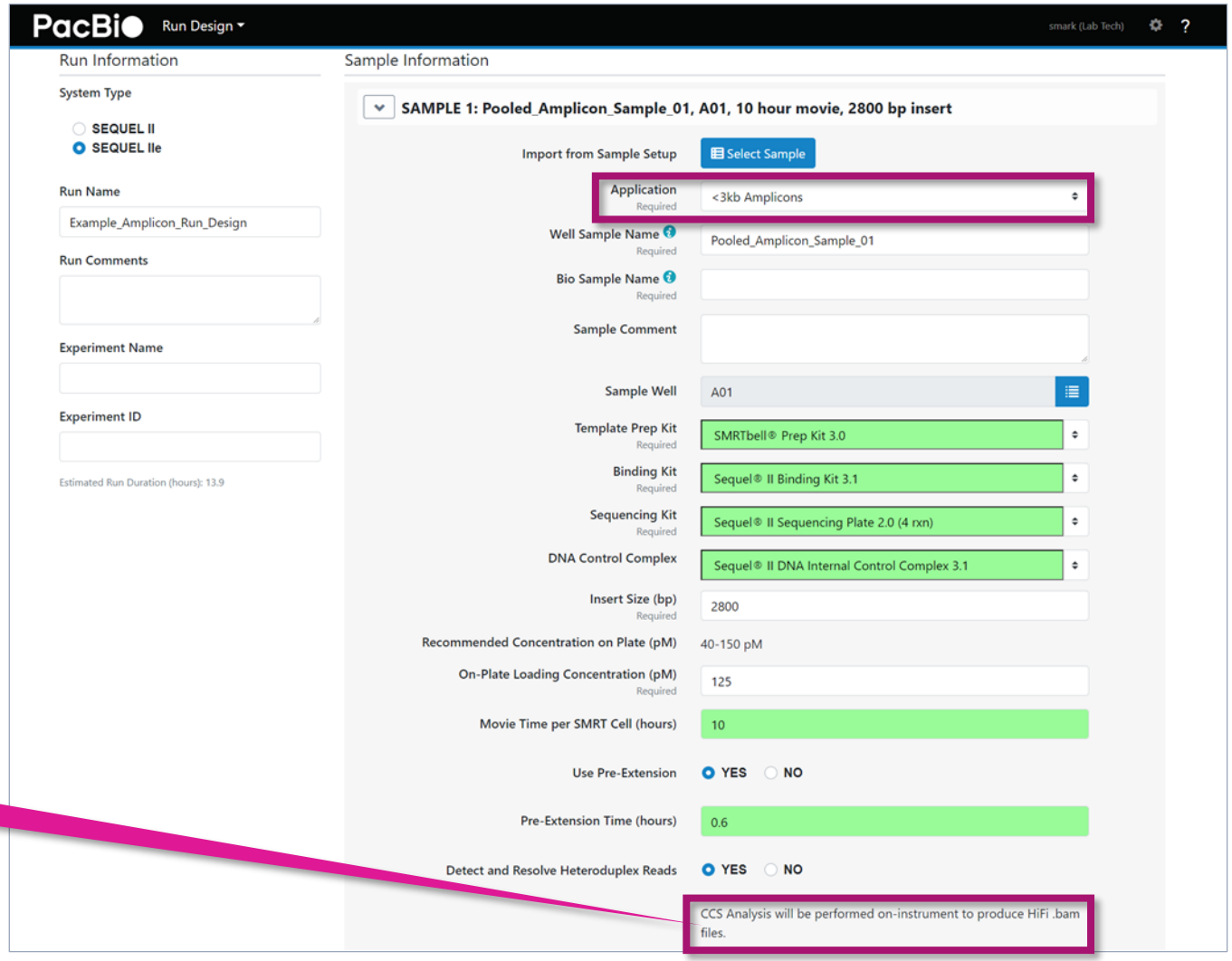
**Example Sample Setup HT mode worksheet for a batch consisting of two amplicon samples (where each sample is comprised of a pooled library containing barcoded amplicons).**

# Amplicon library Run Design guidance

## Follow SMRT Link Run Design instructions to set up a sequencing run using recommended settings for amplicon samples

- Select <3 kb Amplicons or ≥3 kb Amplicons from the Application field drop-down menu in SMRT Link Run Design
- The following fields are auto-populated and high-lighted in green:
  - Template Prep Kit
  - Binding Kit
  - Sequencing Kit
  - DNA Control Complex
  - Movie Time Per SMRT Cell
  - Pre-Extension Time

**Note:** By default, all newly created run designs (regardless of application type) will specify to **automatically** perform CCS analysis and output **only** HiFi reads

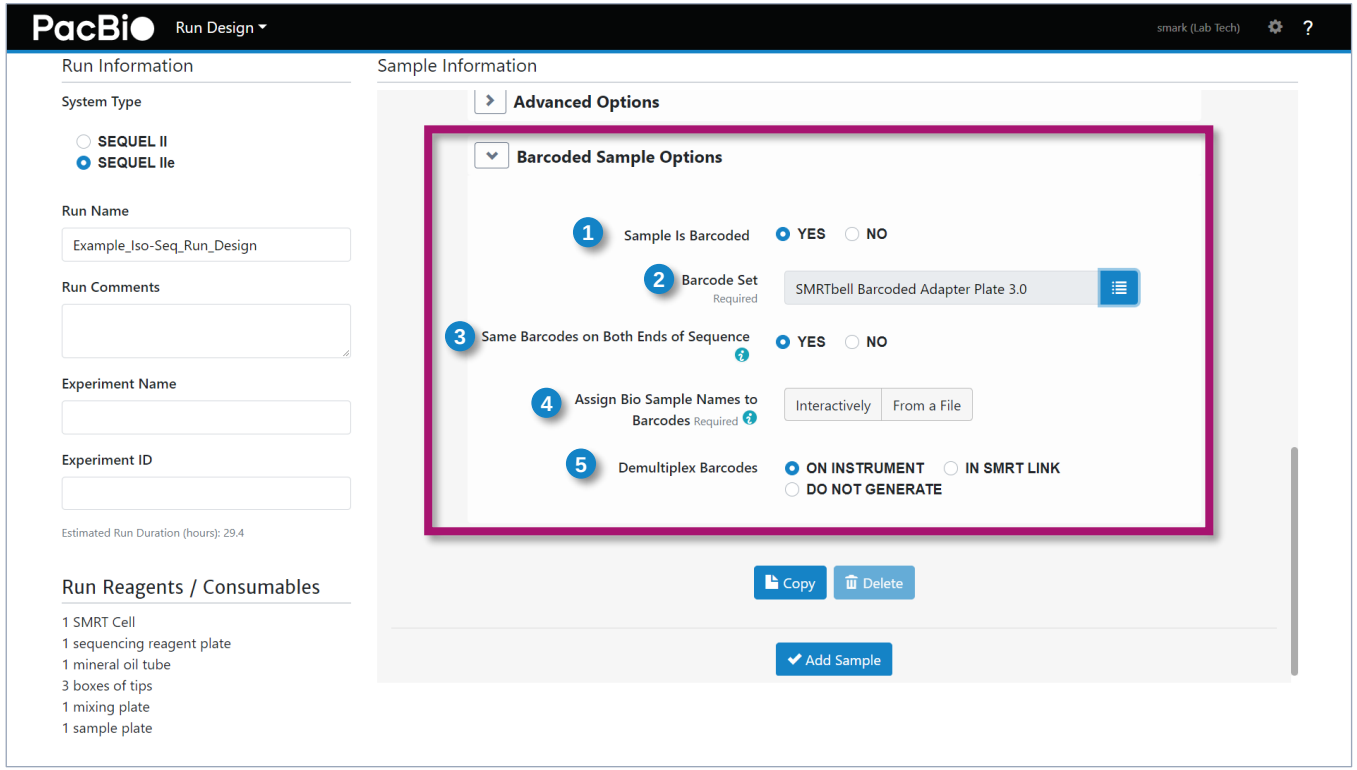



**Example sample information entered into Run Design for sequencing a <3 kb amplicon sample.**

# Amplicon library Run Design guidance (cont.)

## OPTIONAL: Run Design setup procedure for automated demultiplexing of pooled amplicon library samples barcoded with barcoded primers, SMRTbell barcoded adapter plate 3.0 or barcoded M13 primers

1. Sample is Barcoded: YES
2. Barcode Set:*
   - Select 'Sequel_RSII_384_barcodes_v1' if samples were barcoded using gene-specific barcoded primers; or
   - Select 'SMRTbell Barcoded Adapter Plate 3.0 (bc2001-bc2096)' if samples were barcoded using SMRTbell barcoded adapter plate 3.0 (102-009-200); or
   - Select 'Barcoded M13 Primer Plate' if samples were barcoded using Barcoded M13 primer plate (102-135-500)
3. Same Barcodes on Both Ends of Sequence:
   - Select 'YES' if the barcode sequences at both ends of the amplicon insert are the same; or
   - Select 'NO' if the barcode sequences are different on each end of the insert.
4. Assign a Biological Sample Name to each barcoded sample using one of two ways: From a (CSV) File or Interactively
5. Specify if barcode demultiplexing is to be performed on-instrument (Sequel IIe system only) or in SMRT Link. (Optionally specify Do Not Generate.)

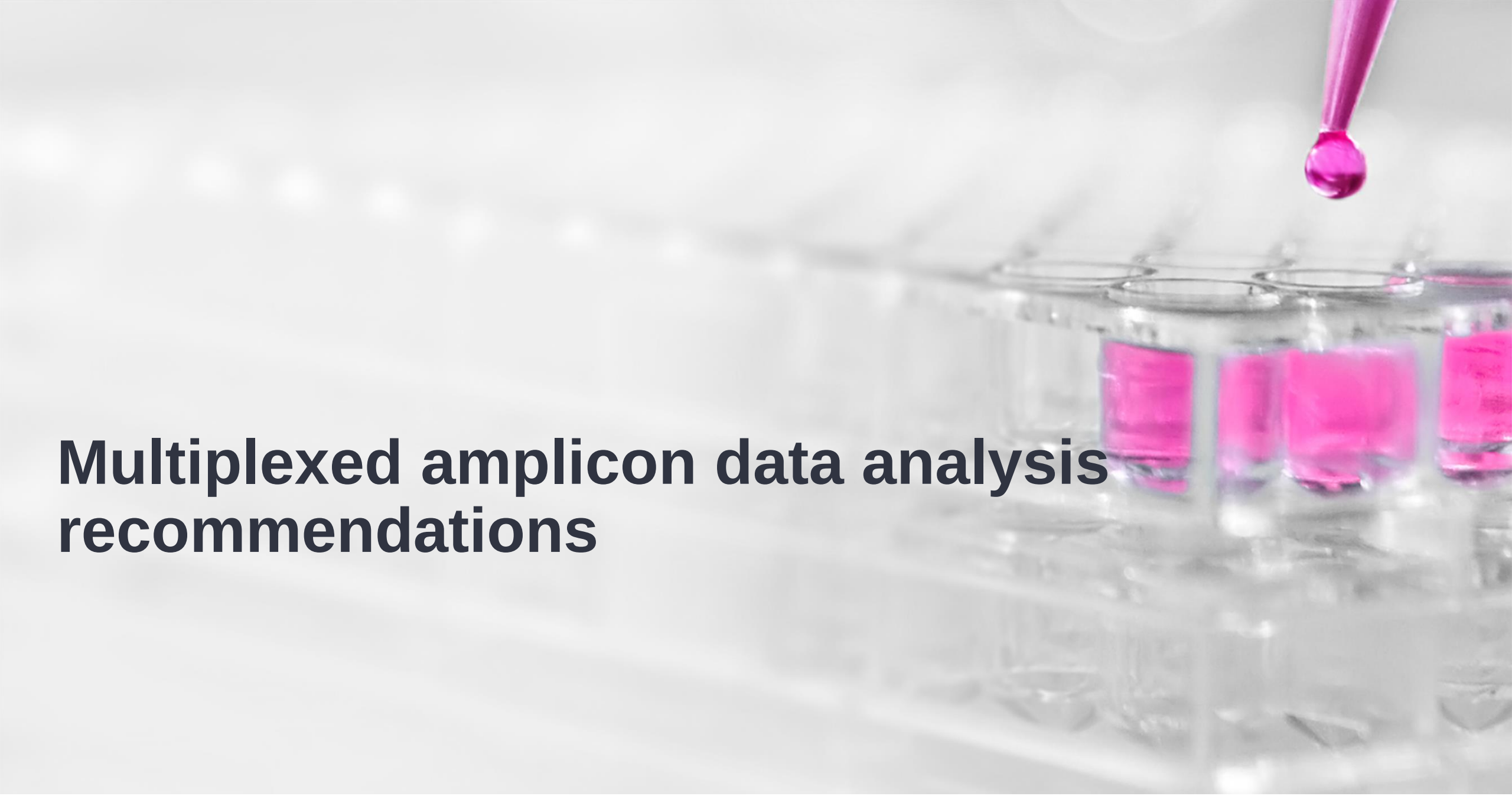



**Example barcoding information entered into Run Design for sequencing a pooled amplicon sample barcoded with SMRTbell barcoded adapter plate 3.0.**



* Refer to our Multiplexing website for application-specific recommendations for barcodes, procedures, ordering sheets and FASTA files required for sample demultiplexing.

# Multiplexed amplicon data analysis recommendations



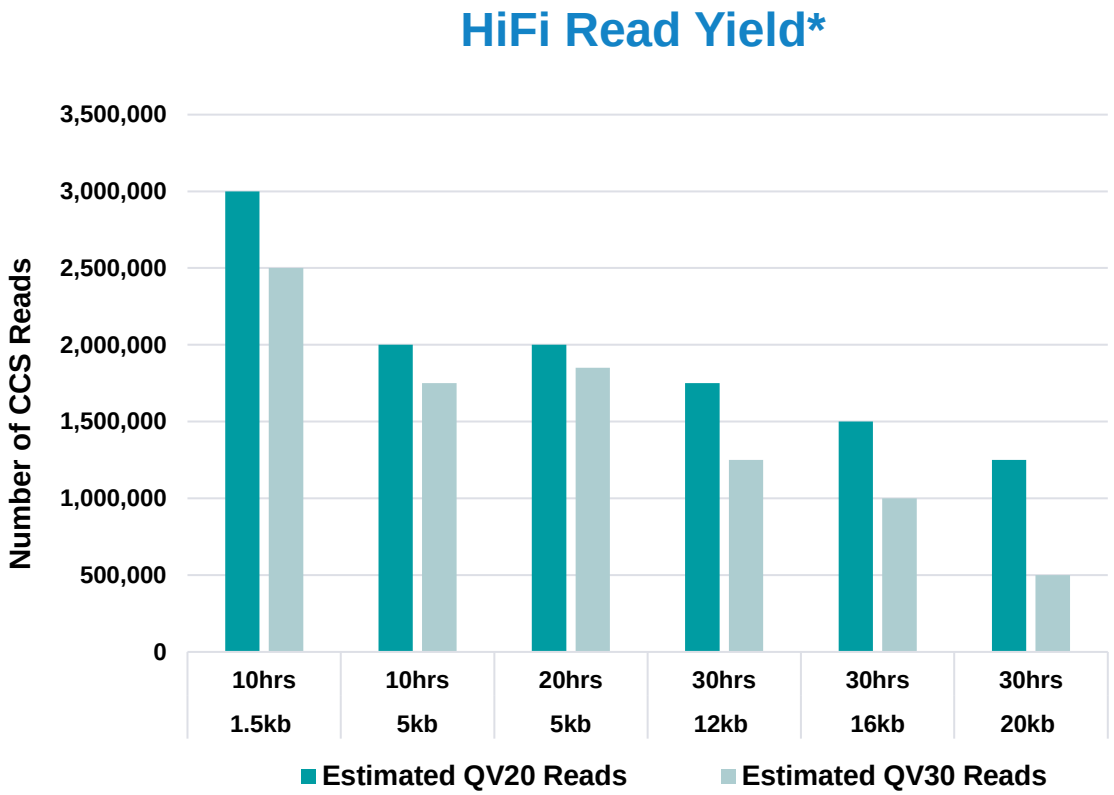
# Multiplexed amplicon data analysis general recommendations

Use SMRT Link and other analysis tools to fully characterize genetic complexity – structural variation, rare SNPs, indels, CNV, microsatellites, haplotypes, and phasing

- Perform circular consensus sequencing (CCS) analysis on-instrument (Sequel IIe system only) or utilize SMRT Link to generate highly accurate (≥Q20) single-molecule long reads (HiFi reads)
- **≥50-fold HiFi read coverage per target locus** is recommended for variant detection applications
- **6,000-fold HiFi read coverage per target locus** is recommended for minor variant detection (1% sensitivity) applications
- Can use SMRT Link to detect, quantitate, and phase single nucleotide polymorphisms within coding regions using the Minor Variants Analysis (MVA) application
- Perform reference-free analysis for complex loci like HLA using pb amplicon analysis
- Output data in standard file formats, (BAM and FASTA/Q) for seamless integration with downstream analysis tools
- HiFi reads are compatible with standard analysis tools for variant calling such as Google DeepVariant

# HiFi read yield performance for different library insert sizes

Generate up to 3 Million HiFi reads or more with the Sequel II and IIe systems depending on your amplicon library size range




| Insert size | Movie length | Estimated Q20 reads | Estimated Q30 reads |
|---|---|---|---|
| 1.5 kb | 10 hrs | 3,000,000 | 2,500,000 |
| 5 kb | 10 hrs | 2,000,000 | 1,750,000 |
| 5 kb | 20 hrs | 2,000,000 | 1,850,000 |
| 12 kb | 30 hrs | 1,750,000 | 1,250,000 |
| 16 kb | 30 hrs | 1,500,000 | 1,000,000 |
| 20 kb | 30 hrs | 1,250,000 | 500,000 |

**Data shown above are for SMRTbell library samples sequenced on a Sequel II system using different movie collection times. Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.**

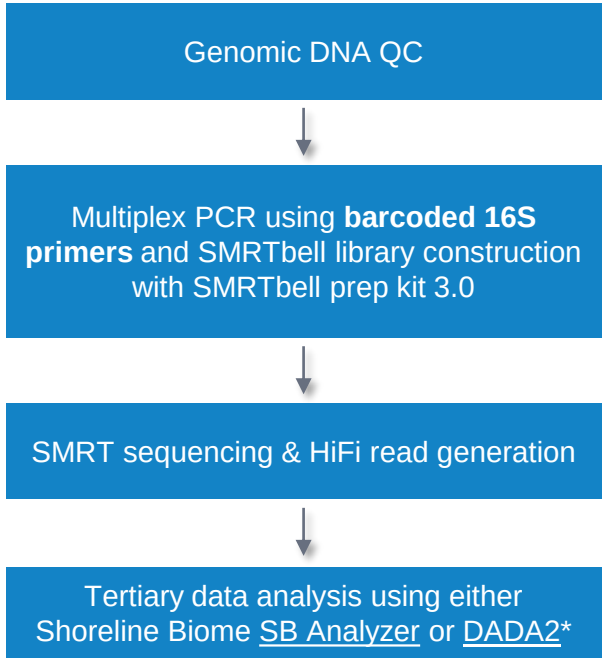
PacBio

* 50% P1 Loading

# Sample preparation recommendations for full-length 16S amplicon sequencing

# SMRTbell library preparation workflow overview for 16S amplicon samples generated with barcoded gene-specific primers

## Follow PacBio's 16S rRNA gene amplification protocol with recommended barcoded primers for generating multiplexed, full-length 16S samples for SMRT sequencing on Sequel II and IIe systems




**Workflow for amplification and sequencing of full-length 16S amplicons generated with barcoded gene-specific PCR primers.**

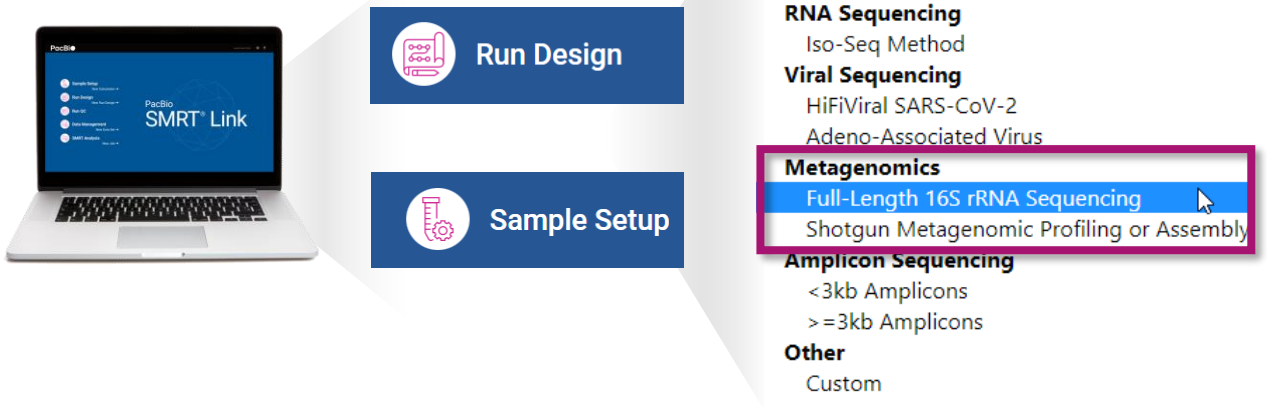
- To reduce cost per sample, up to 192 barcoded 16S amplicon samples may be pooled for SMRTbell library construction and sequencing on a single SMRT Cell 8M by using the procedure below:
  - Procedure & checklist – Multiplexed amplicon library preparation using SMRTbell Prep Kit 3.0 (102-359-000) (follow workflow instructions for processing primer barcoded samples)
- For guidance on recommended gene-specific primer sequences and PCR conditions for amplification of full-length 16S genes (V1-V9 regions) from bacterial DNA isolated from metagenomic samples, refer to the procedure below:
  - Procedure & checklist – Amplification of bacterial full-length 16S rRNA gene with barcoded primers (101-599-700)
  - **NOTE:** For carrying out SMRTbell library construction using pooled, barcoded 16S amplicon DNA input, follow workflow instructions in Procedure & checklist 102-359-000 for processing primer barcoded samples
- Tertiary analysis of HiFi reads can be performed using either Shoreline Biome SB Analyzer software or with the DADA2* analysis pipeline.



* Callahan, B. J., et al. (2019) High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. Nucleic Acids Research, 47(18), e103

# Sample Setup & Run Design recommendations for 16S amplicon libraries

## In SMRT Link Sample Setup & Run Design, select ‘Metagenomics’ and choose ‘Full-Length 16S rRNA Sequencing’ for application type





**Sequel II binding kit 3.1 & cleanup beads (102-333-400) is recommended for preparing 16S amplicon samples for sequencing.**

**Sequel II binding kit 3.1 & cleanup beads** (102-333-300) includes:

- Sequencing primer 3.1
- Sequel II polymerase 2.1
- DNA internal control 3.1 (defined 2 kb template bound to Polymerase 2.1)
- SMRTbell cleanup beads for complex cleanup

- After specifying your application type, **SMRT Link** auto-fills selected Sample Setup and Run Design parameter fields with default recommended values*

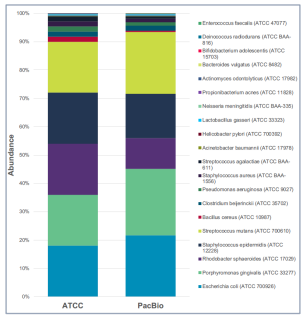
| Amplicon library type | Recommended binding kit |
|---|---|
| 16S | Sequel II binding kit 3.1 |

* See **Quick reference card – loading and pre-extension recommendations for Sequel II and IIe systems** (101-769-100) for any updates to sample setup and run design recommendations for specific applications.

# 16S data analysis workflow recommendations

## PacBio recommends using GenDx's NGSengine software for HLA typing

1. Generate HiFi reads
2. Demultiplex barcodes
3. Tertiary data analysis using either Shoreline Biome SB Analyzer or DADA2*




1. Perform CCS analysis on-instrument (Sequel IIe system only) or in SMRT Link to generate highly accurate (≥Q20) single-molecule long reads (HiFi reads)

2. Demultiplex barcodes on-instrument (Sequel IIe system only) or in SMRT Link to separate HiFi reads by sample barcode
   - Barcode FASTA files for demultiplexing can be downloaded from PacBio's Multiplexing website

3. Analyze 16S data using Shoreline Biome SB Analyzer or DADA2

DADA2 Amplicon Sequencing. Exactly. Version 1.8

- Open-source
- Well documented
- R package
- Easy and fast

An example HiFi read data set for a MSA-1003 mock community sample is available for download from PacBio (Link)

# Example workflow: 192-plex 16S amplicon library preparation using barcoded gene-specific primers

## MSA-1003 Mock Community Sample Description

- MSA-1003 is a controlled, pre-defined, standardized reference material that can help with metagenomic analysis protocol development optimization, verification, and quality control
- 20 Strain Staggered Mix Genomic Material (ATCC MSA-1003) https://www.atcc.org/products/all/MSA-1003.aspx
- MSA-1003 sample is a mock microbial community that mimics mixed metagenomic samples
- MSA-1003 sample comprises genomic DNA prepared from fully sequenced, characterized, and authenticated ATCC Genuine Cultures that were selected by ATCC based on relevant phenotypic and genotypic attributes, such as Gram stain, GC content, genome size, and spore formation
- For the example data shown in this presentation, replicate MSA-1003 samples were processed in parallel to generate a 192-plex pooled 16S SMRTbell library using barcoded gene-specific primers and SMRTbell express template prep kit 2.0

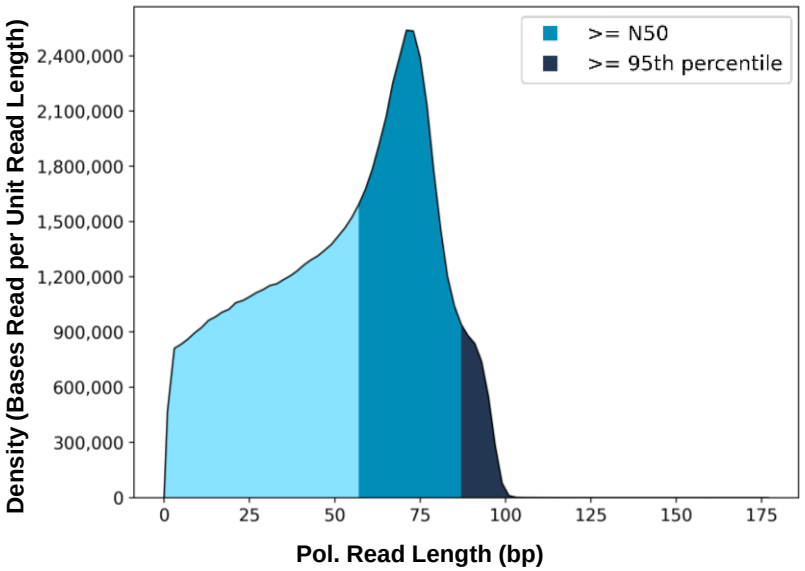
| % | MSA-1003 component |
|---|---|
| 0.18 | *Acinetobacter baumannii* (ATCC 17978) |
| 1.80 | *Bacillus cereus* (ATCC 10987) |
| 0.02 | *Bacteroides vulgatus* (ATCC 8482) |
| 0.02 | *Bifidobacterium adolescentis* (ATCC 15703) |
| 1.80 | *Clostridium beijerinckii* (ATCC 35702) |
| 0.18 | *Cutibacterium acnes* (ATCC 11828) |
| 0.02 | *Deinococcus radiodurans* (ATCC BAA-816) |
| 0.02 | *Enterococcus faecalis* (ATCC 47077) |
| 18.0 | *Escherichia coli* (ATCC 700926) |
| 0.18 | *Helicobacter pylori* (ATCC 700392) |
| 0.18 | *Lactobacillus gasseri* (ATCC 33323) |
| 0.18 | *Neisseria meningitidis* (ATCC BAA-335) |
| 18.0 | *Porphyromonas gingivalis* (ATCC 33277) |
| 1.80 | *Pseudomonas aeruginosa* (ATCC 9027) |
| 18.0 | *Rhodobacter sphaeroides* (ATCC 17029) |
| 0.02 | *Schaalia odontolytica* (ATCC 17982) |
| 1.80 | *Staphylococcus aureus* (ATCC BAA-1556) |
| 18.0 | *Staphylococcus epidermidis* (ATCC 12228) |
| 1.80 | *Streptococcus agalactiae* (ATCC BAA-611) |
| 18.0 | *Streptococcus mutans* (ATCC 700610) |

https://www.atcc.org/products/all/MSA-1003.aspx

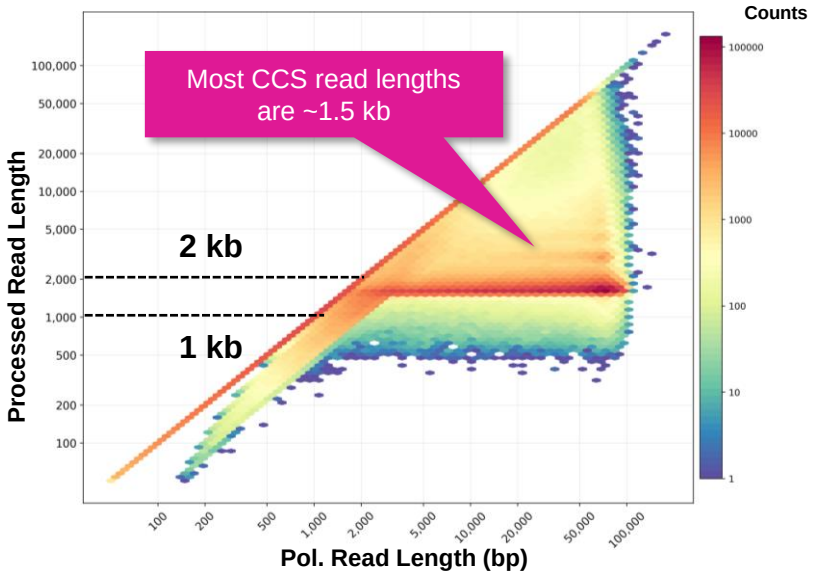
# Example workflow: 192-plex 16S amplicon library preparation using barcoded gene-specific primers (cont.)

## Primary sequencing metrics

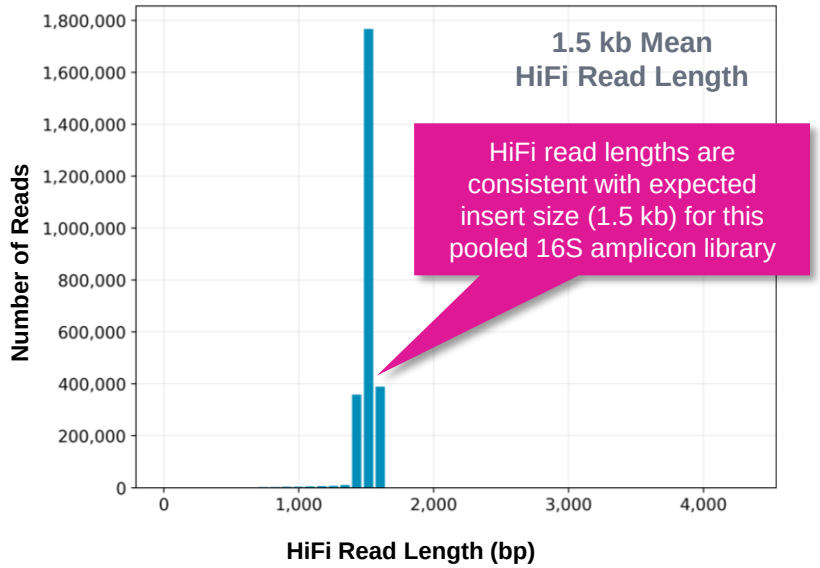
### Base Yield Density




### Read Length Density




### HiFi Read Length Distribution




- 192-plex pooled 16S amplicon library (1.5 kb) constructed with barcoded 16S primers
- Run on a Sequel II system with Polymerase 2.1 / 50 pM on-plate concentration / 10-hr movie time / 1-hr Pre-extension time

| | |
|---|---|
| **Raw Base Yield** | 128.5 Gb |
| **Mean Polymerase Read Length** | 24.5 kb |
| **P0** | 29.7% |
| **P1** | 65.3% |
| **P2** | 5.0% |

| | |
|---|---|
| **HiFi Reads** | 2.6 M |
| **HiFi Base Yield** | 3.9 Gb |
| **Mean HiFi Read Length** | 1,532 bp |
| **Median HiFi Read Quality** | Q41 |
| **HiFi Read Mean # of Passes** | 18 |



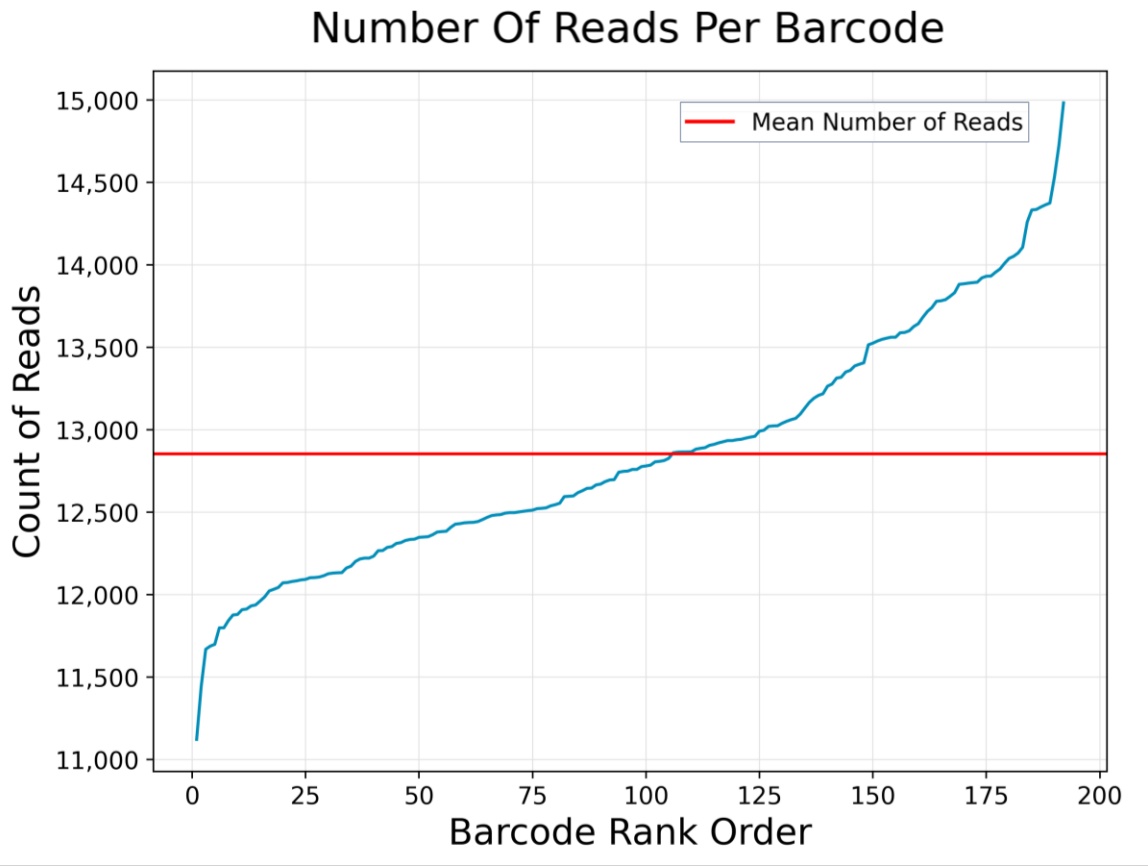
Read lengths, reads/data per SMRT Cell and other sequencing performance results vary based on sample quality/type and insert size.

# Example workflow: 192-plex 16S amplicon library preparation using barcoded gene-specific primers (cont.)

## Barcode demultiplexing results

**192-plex 16S SMRTbell library CCS analysis and barcode demultiplexing results**

| Analysis | Analysis metric | Value |
|---|---|---|
| CCS | HiFi (≥Q20 CCS) reads | 2,568,971 |
| Demultiplex barcodes | Unique barcodes detected | 192 |
| | Total barcoded HiFi reads | 2,468,174 |
| | Barcode recovery rate | 96% |
| | Mean barcoded HiFi reads per sample | 12,855 |
| | Max. barcoded HiFi reads per sample | 14,983 |
| | Min. barcoded HiFi reads per sample | 11,121 |
| | Mean barcoded HiFi read length | 1,495 |



# Example workflow:  192-plex 16S amplicon library preparation using barcoded gene-specific primers (cont.)

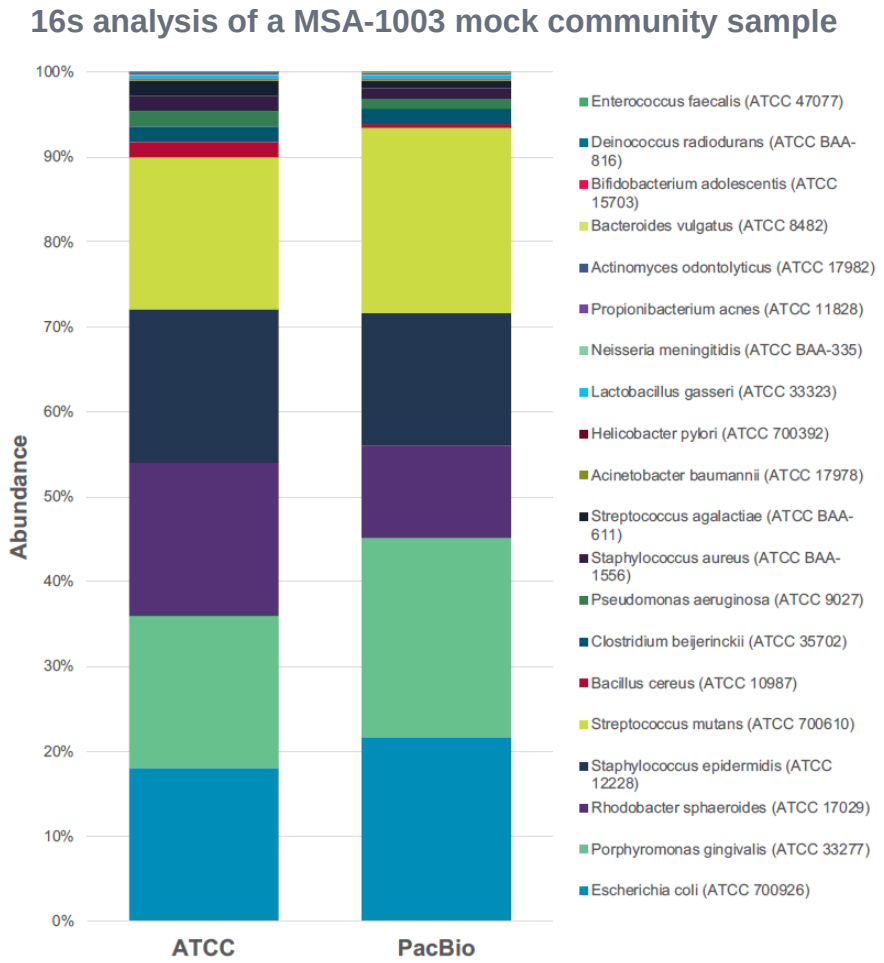
## Example taxonomic classification results for 192-plex 16S library sample

**MSA-1003 sample description**

20 Strain staggered mix genomic material (ATCC MSA-1003)
https://www.atcc.org/products/all/MSA-1003.aspx

16S HiFi sequencing data set reproduces the expected composition of the MSA-1003 mock community  sample

Download and explore this 16S HiFi dataset further

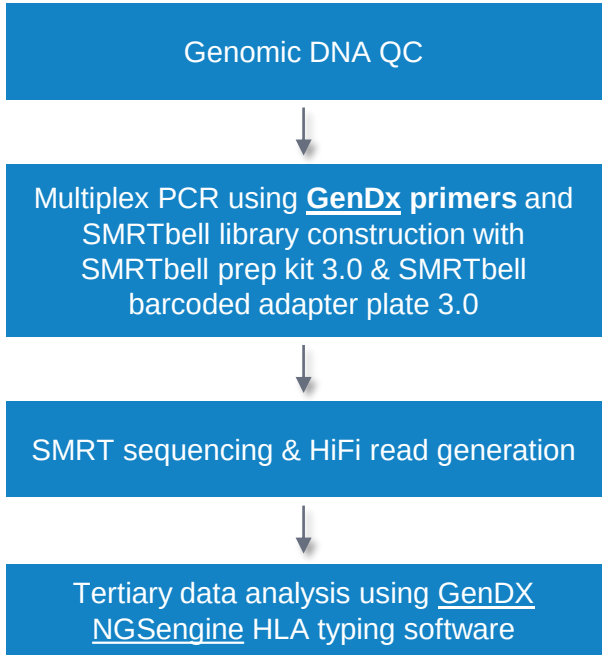



**Full-length (V1-V9) 16S amplicon samples were pooled at 192-plex and sequenced on a single SMRT Cell 8M. PacBio results shown in bar graph reflect the average abundance values derived from the pooled MSA-1003 replicate samples.**

# Sample preparation recommendations for human leukocyte antigen (HLA) amplicon sequencing

# SMRTbell library preparation workflow overview for HLA amplicon samples generated with GenDx HLA typing kits

## GenDx (gendx.com) offers validated HLA primers with ready-to-use PCR master mixes, protocols and tools for SMRT sequencing on Sequel II and IIe systems



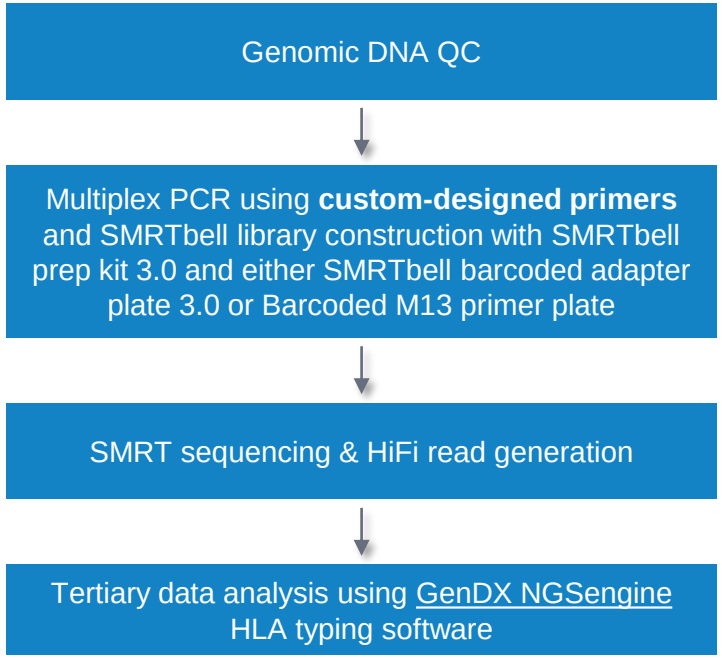

**Workflow for amplification and sequencing of HLA amplicons generated with GenDx PCR primers.**

- Depending on your project requirements, 3 kits are available from GenDX for typing up to 11 HLA loci
- PacBio has validated the GenDx NGSgo-MX6-1 Kit
  - Features six HLA loci in one tube for the amplification of HLA-A, -B, -C, -DRB1, -DQB1 and –DPB1 (3 – 6 kb target amplicon sizes)
  - Kit consists of 3 reagent tubes supporting up to 96 PCR reactions total:
    - NGSgo-MX6-1 primer mix
    - GenDx-LongMix PCR master mix (4x)
    - Nuclease-free water


- To reduce cost per sample, multiple samples may be barcoded using SMRTbell barcoded adapter plate 3.0 (102-009-200) and pooled for sequencing on a single SMRT Cell
- Up to 96 samples may be pooled for sequencing on the Sequel II and IIe systems by using the procedure below to barcode samples using barcoded adapters:
  - Procedure & dhecklist – Multiplexed amplicon library preparation using SMRTbell Prep Kit 3.0 (102-359-000)
- HLA typing analysis of HiFi reads can be performed with GenDx NGSengine software

# SMRTbell library preparation workflow overview for HLA amplicon samples generated with customer-designed assays

## HLA amplicon samples generated with custom-designed assays can be constructed to SMRTbell libraries and sequenced on Sequel II and IIe Systems



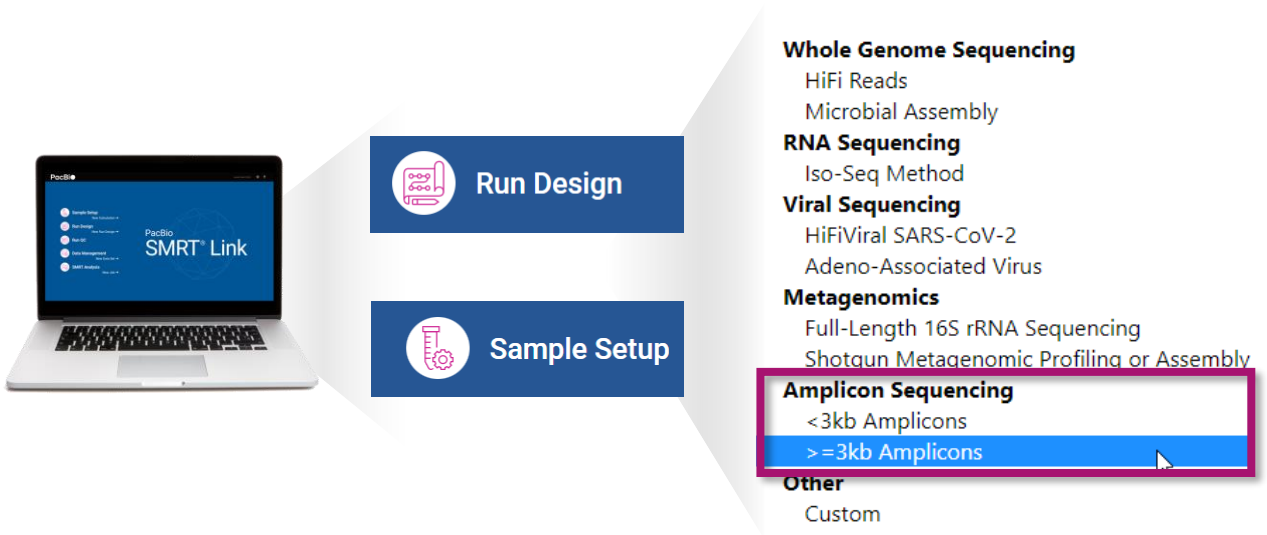

**Workflow for amplification and sequencing of HLA amplicons generated with custom-designed PCR primers.**

- **NOTE:** Users may design their own custom primers to amplify HLA gene targets for SMRT Sequencing – but be aware that painstaking and time-consuming validation experiments may be required to develop an optimized sample preparation workflow.
- We therefore highly recommend using validated and ready-to-use primers from GenDx to amplify HLA gene targets for PacBio sequencing

- To reduce cost per sample, multiple samples may be barcoded and pooled for sequencing on a single SMRT Cell
- Up to 96 barcoded samples may be pooled for sequencing on the Sequel II and IIe systems using either of two available barcoding strategies:
  - Use barcoded adapters for barcoding by following Procedure & checklist – Preparing multiplexed amplicon libraries using SMRTbell prep kit 3.0 (102-359-000); or
  - Use barcoded M13 primers for barcoding by following Procedure & checklist – Preparing multiplexed amplicon libraries using PacBio barcoded M13 primers and SMRTbell prep kit 3.0 (101-921-300)
- HLA typing analysis of HiFi reads can be performed with GenDx NGSengine software

# Sample Setup & Run Design recommendations for HLA amplicon libraries

## In SMRT Link Sample Setup & Run Design, select ‘Amplicon Sequencing’ and choose ‘≥3 kb Amplicons’ for application type





**Sequel II binding kit 3.2 & cleanup beads (102-333-300) is recommended for preparing HLA amplicon samples for sequencing.**

**Sequel II binding kit 3.2 & cleanup beads** (102-333-300) includes:

- Sequencing primer 3.2
- Sequel II polymerase 2.2
- DNA internal control 3.2 (defined 11 kb template bound to Polymerase 2.2)
- SMRTbell cleanup beads for complex cleanup

- After specifying your application type, **SMRT Link** auto-fills selected Sample Setup and Run Design parameter fields with default recommended values*

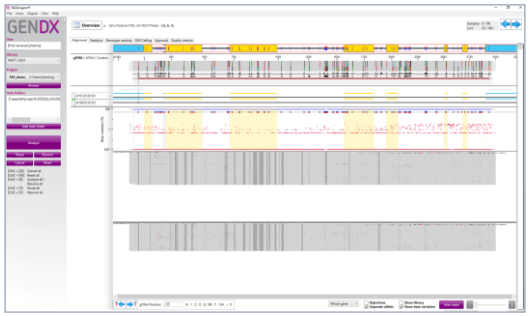
| Amplicon library type | Recommended binding kit |
|---|---|
| HLA | Sequel II binding kit 3.2 |

* See **Quick reference card – loading and pre-extension recommendations for Sequel II and IIe systems** (101-769-100) for any updates to sample setup and run design recommendations for specific applications.

# HLA data analysis workflow recommendations

## PacBio recommends using GenDx's NGSengine software for HLA typing

1. Generate HiFi reads
2. Demultiplex barcodes
3. GENdx NGSengine HLA typing software




1. Perform CCS analysis on-instrument (Sequel IIe system only) or in SMRT Link to generate highly accurate (≥Q20) single-molecule long reads (HiFi reads)
   - In SMRT Link, recommend setting Minimum CCS Read Length = 3000 bp

2. Demultiplex barcodes on-instrument (Sequel IIe system only) or in SMRT Link to separate HiFi reads by sample barcode
   - Barcode FASTA files for demultiplexing can be downloaded from PacBio's Multiplexing website

3. Utilize GenDX NGSengine software to perform HLA Typing analysis of de-multiplexed HiFi data
   - Recommended Input: **≥100 HiFi reads per HLA locus**
   - NGSengine accepts as input batches of demultiplexed FASTQ files containing HiFi sequencing data for 1 or more pooled HLA loci.
   - HLA locus assignment and HLA typing for each demultiplexed sample dataset can be performed on a Windows-based laptop computer.

An example demultiplexed HLA dataset analyzed using NGSengine is available for download from PacBio (Link)

# Example workflow:  96-plex HLA amplicon library preparation using GenDx NGSgo-MX6-1 kit and barcoded adapters

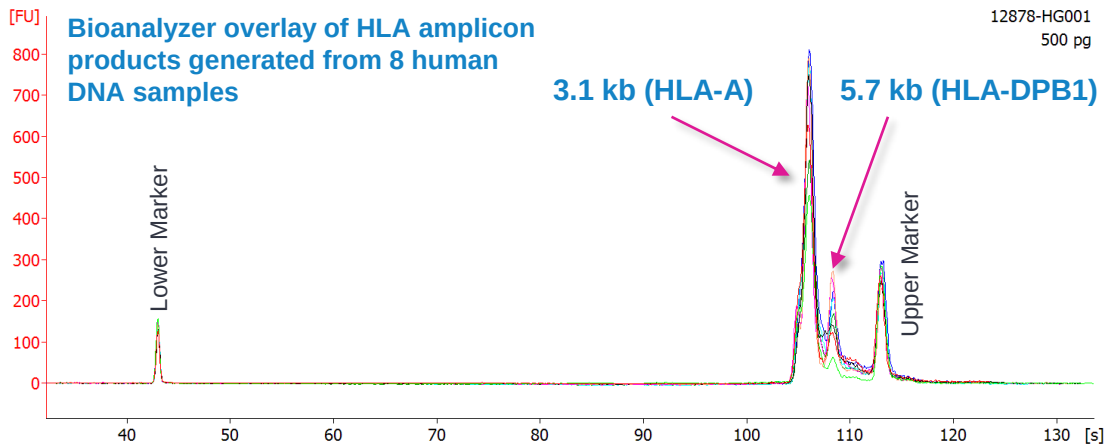
## HLA amplicon generation using GenDX NGSgo-MX6-1 kit

**PCR product yields for 8 human gDNA samples amplified using the GenDx NGSgo-MX6-1 kit**

| Sample ID | Nist id | # Of replicate PCR reactions* | Purified PCR product conc. (ng/µL) | Purified PCR product volume (µL) | Total mass of purified PCR products (ng) | Mass of purified PCR product per reaction (ng) |
|---|---|---|---|---|---|---|
| **NA12878** | HG001 | 4 | 228 | 16 | 3,648 | 912 |
| **NA24385** | HG002 | 4 | 250 | 16 | 4,000 | 1,000 |
| **NA24149** | HG003 | 4 | 240 | 16 | 3,840 | 960 |
| **NA24143** | HG004 | 4 | 221 | 16 | 3,392 | 848 |
| **NA24631** | HG005 | 4 | 250 | 16 | 4,000 | 1,000 |
| **NA24695** | HG007 | 4 | 252 | 16 | 4,032 | 1,008 |
| **NA06896** | N/A | 4 | 254 | 16 | 4,064 | 1,016 |
| **C1-218** | N/A | 4 | 230 | 16 | 3,680 | 912 |

* For each PCR reaction, 200 ng of input gDNA per sample was used.  The fragment size of the input genomic DNA samples ranged from 22 kb to 147 kb (mode).

- Four (4) replicate PCR reactions were performed for each human gDNA sample (input gDNA size mode ~22 kb – 147 kb)
- For each sample, replicate PCR reaction products were pooled
- Bioanalyzer sizing QC results are consistent with expected range of PCR amplicon sizes (~3.1 – 5.7 kb) for these HLA samples



# Example workflow:  96-plex HLA amplicon library preparation using GenDx NGSgo-MX6-1 kit and barcoded adapters (cont.)

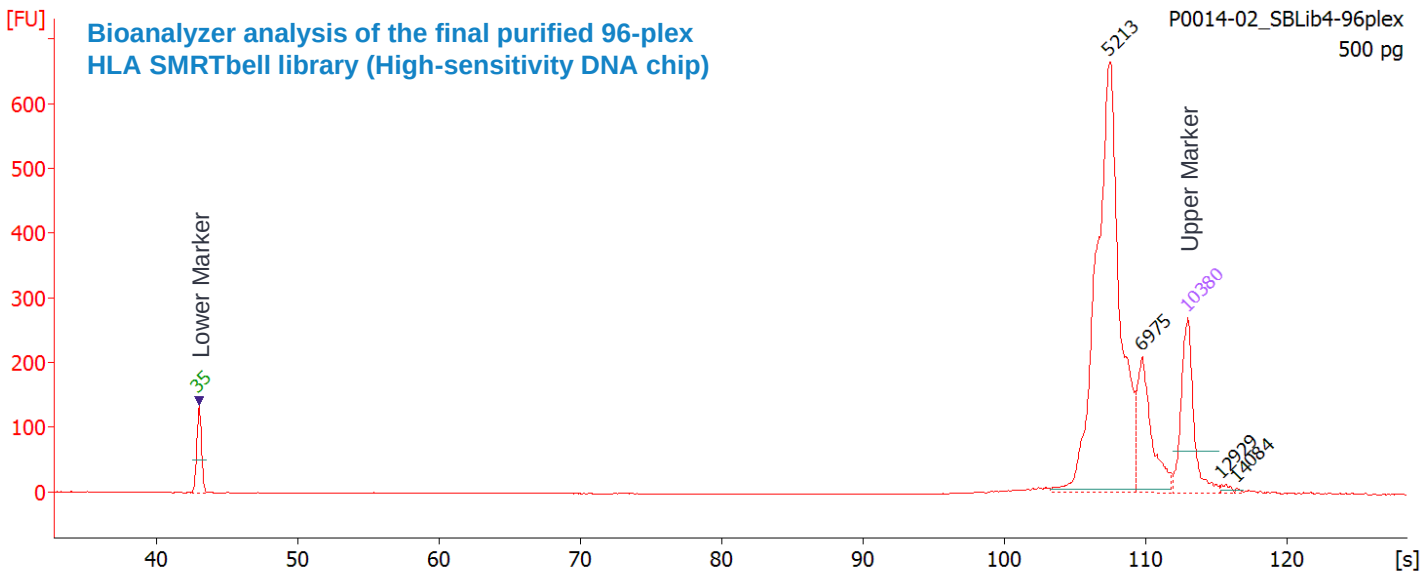
## HLA amplicon SMRTbell library construction using SMRTbell express template prep kit 2.0

**96-plex HLA SMRTbell library construction yield**

| SMRTbell library | Number of pooled samples | Input mass of PCR DNA per sample for library construction | Total input mass | Purified library conc. (ng/µL)* | Purified library volume (µL) | Purified library mass (ng) | Library construction yield (%) |
|---|---|---|---|---|---|---|---|
| 96-Plex HLA Library | 96 | 150 | 14,400 | 135 | 96 | 12,960 | 90 |

* The final 96-plex HLA SMRTbell express TPK 2.0 library was purified using two rounds of 0.6X AMPure PB purification at the end of the procedure

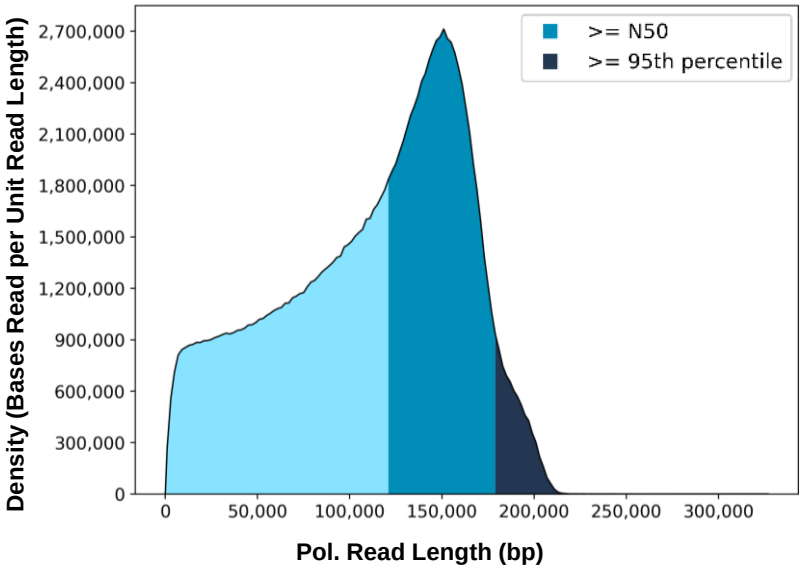
- Each of eight (8) human gDNA samples that was amplified using the GenDX NGSgo-MX6-1 kit was barcoded with 12 unique barcoded adapters to generate a 96-plex pooled HLA SMRTbell library
- Mean size of the final purified 96-plex HLA SMRTbell library was ~5700 bp



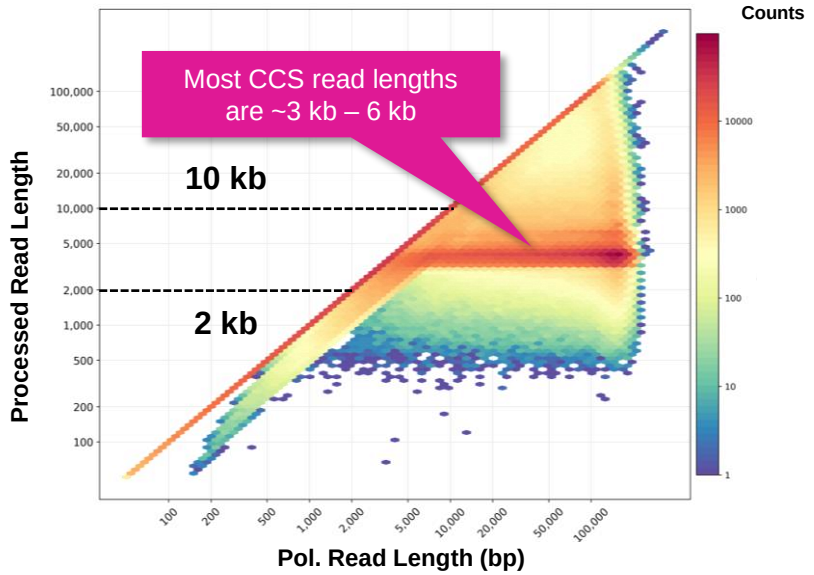
# Example workflow: 96-plex HLA amplicon library preparation using GenDx NGSgo-MX6-1 kit and barcoded adapters (cont.)

## Primary sequencing metrics

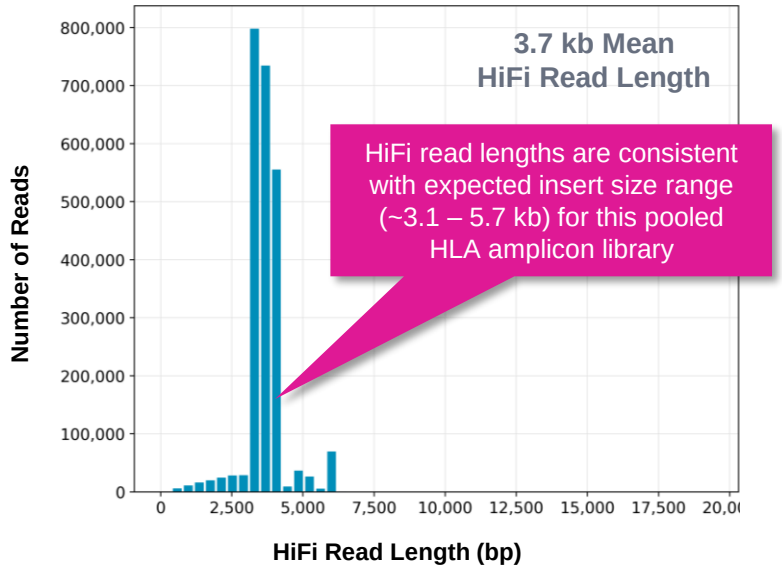
### Base Yield Density




### Read Length Density




### HiFi Read Length Distribution




- 96-plex pooled HLA amplicon library (~3.1 – 5.7 kb) constructed with barcoded adapters
- Run on a Sequel II system with Polymerase 2.0 / 60 pM on-plate concentration / 20-hr movie time / 1.4-hr Pre-extension time

| | |
|---|---|
| **Raw Base Yield** | 276.7 Gb |
| **Mean Polymerase Read Length** | 52.8 kb |
| **P0** | 32.0% |
| **P1** | 65.5% |
| **P2** | 2.6% |

| | |
|---|---|
| **HiFi Reads** | 2.4 M |
| **HiFi Base Yield** | 8.7 Gb |
| **Mean HiFi Read Length** | 3,680 bp |
| **Median HiFi Read Quality** | Q40 |
| **HiFi Read Mean # of Passes** | 19 |



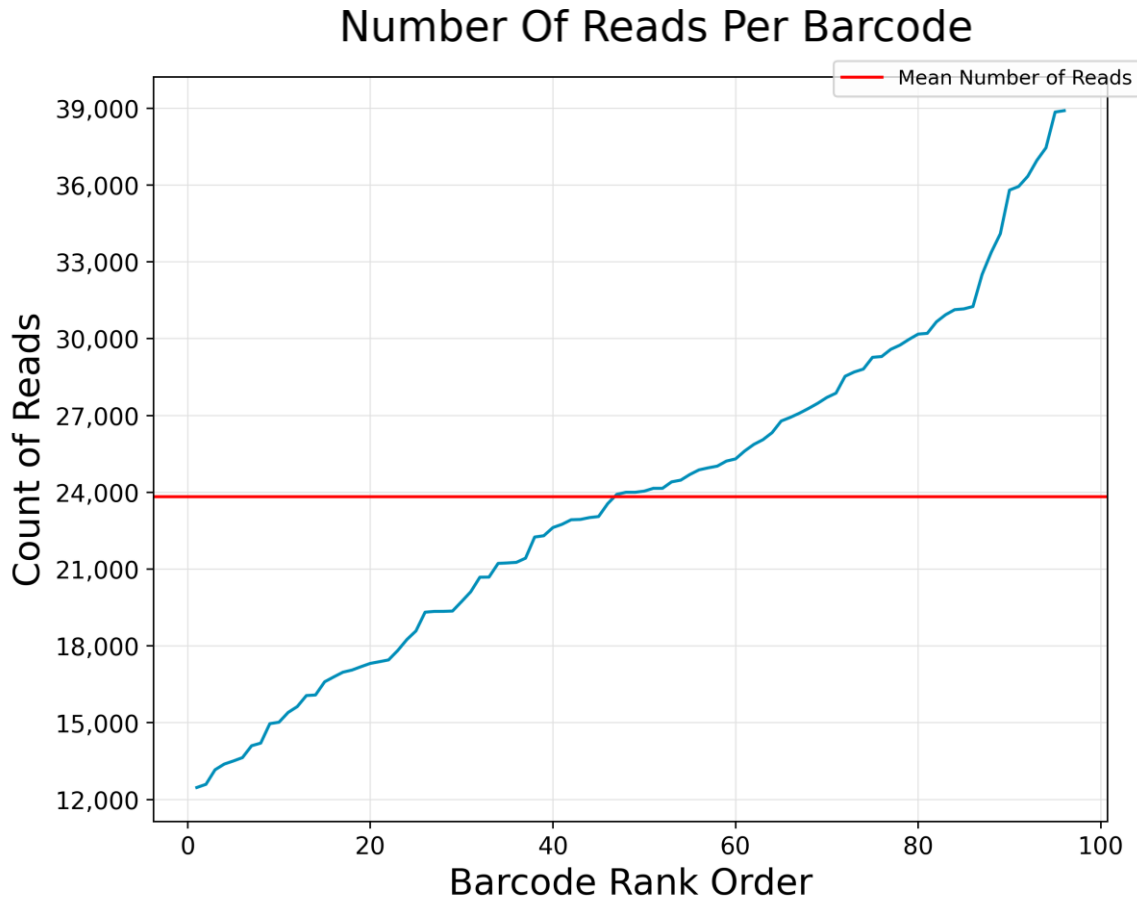
Read lengths, reads/data per SMRT Cell and other sequencing performance results vary based on sample quality/type and insert size.

# Example workflow: 96-plex HLA amplicon library preparation using GenDx NGSgo-MX6-1 kit and barcoded adapters (cont.)

## Barcode demultiplexing results

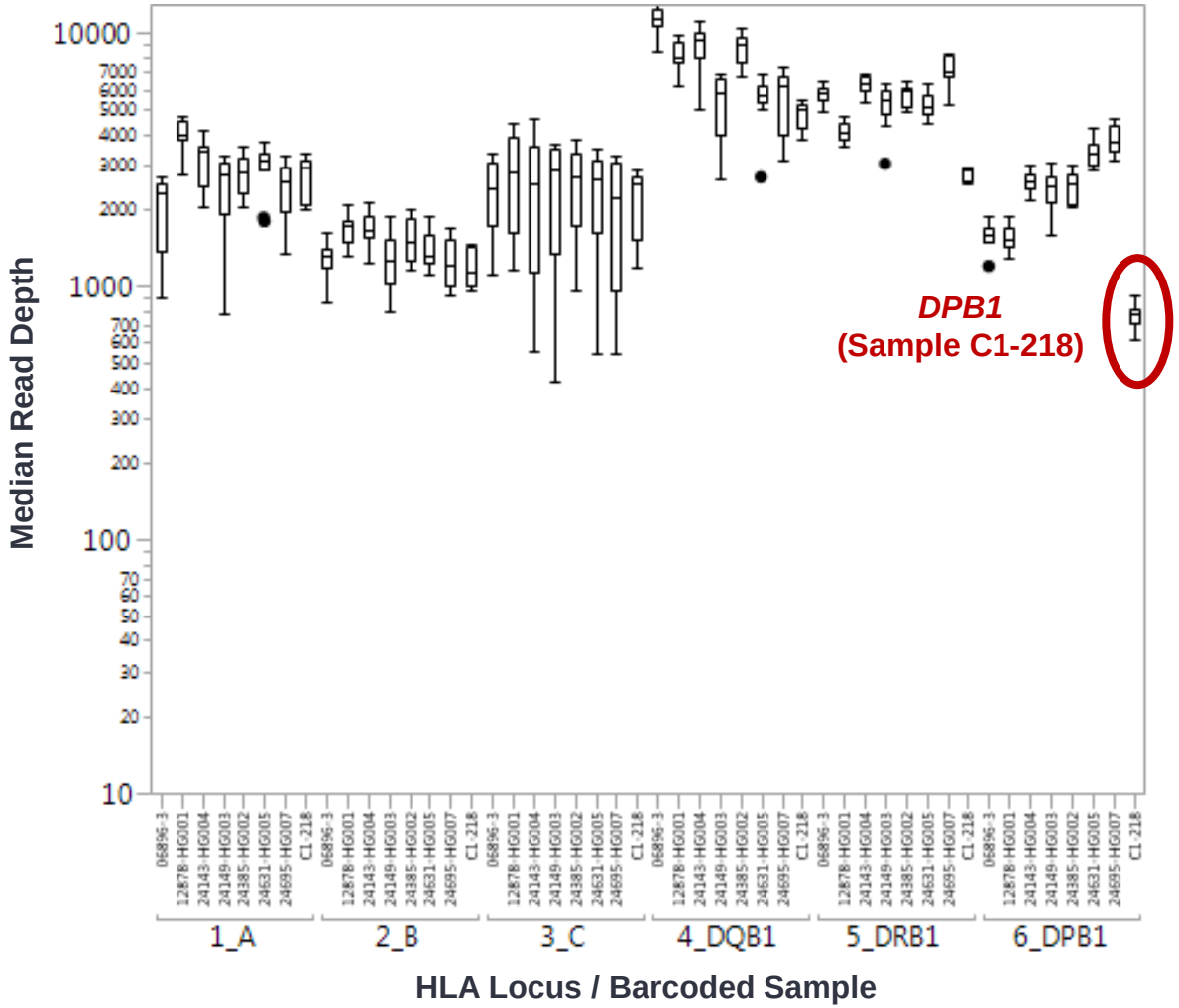
**96-plex HLA SMRTbell library CCS analysis and barcode demultiplexing results**

| Analysis | Analysis metric | Value |
|---|---|---|
| CCS | HiFi (≥Q20 CCS) reads | 2,370,623 |
| Demultiplex barcodes | Unique barcodes detected | 96 |
| | Total barcoded HiFi reads | 2,288,576 |
| | Barcode recovery rate | 97% |
| | Mean barcoded HiFi reads per sample | 23,839 |
| | Max. barcoded HiFi reads per sample | 38,905 |
| | Min. barcoded HiFi reads per sample | 12,469 |
| | Mean barcoded HiFi read length | 3,671 |



# Example workflow: 96-plex HLA amplicon library preparation using GenDx NGSgo-MX6-1 kit and barcoded adapters (cont.)

## HLA read depth analysis using GENdx NGSengine




- HLA locus *DPB1* generated the lowest depth of coverage with sample C1-218
  - **Note:** C1-218 gDNA sample QC showed the highest degree of DNA fragmentation with a size distribution mode of 22 kb, whereas other gDNA samples included in this dataset showed a starting modal size range from ~27 – 147 kb.
  - *DPB1* is the longest HLA amplicon (5.7 kb) amplified using the GenDX NGSgo-MX1 kit.

To help ensure adequate read coverage, particularly for DPB1 alleles, **we recommend pooling a maximum of 96 samples per Sequel II System SMRT Cell 8M** for HLA sequencing

# Example workflow: 96-plex HLA amplicon library preparation using GenDx NGSgo-MX6-1 kit and barcoded adapters (cont.)

## HLA typing results using GENdx NGSengine software

***Sample:*** **demultiplex.24143-HG004**

**Full typing result**

| | *Allele 1* | *Allele 2* | *CWD 1* | *CWD 2* | *Review status* |
|---|---|---|---|---|---|
| **HLA-A** | **01:01:01:01** | **33:01:01:01** | C | C | Not reviewed |
| **HLA-B** | **14:02:01:01** | **35:08:01:01** | C | C | Not reviewed |
| **HLA-C** | **04:01:01:06** | **08:02:01:01** | No | C | Not reviewed |
| **DRB1** | **04:04:01** | **10:01:01:01** | C | C | Not reviewed |
| **DQB1** | **04:02:01:06** | **05:01:01:05** | No | No | Not reviewed |
| **DPB1** | **04:01:01:01** | **04:01:01:01** | C | C | Not reviewed |

**Example HLA typing report for 24143 (HG004) sample analyzed with GenDX NGSengine.**

An example demultiplexed HLA dataset analyzed using NGSengine is available for download from PacBio (Link)

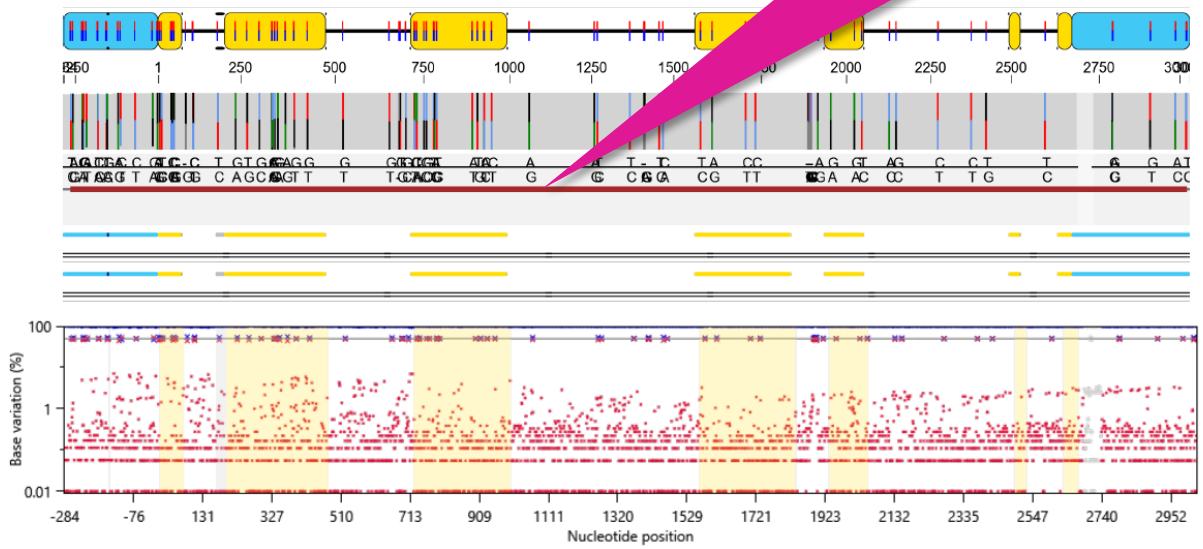
***Sample:*** **demultiplex.24143-HG004**
***Locus:*** **HLA-B**

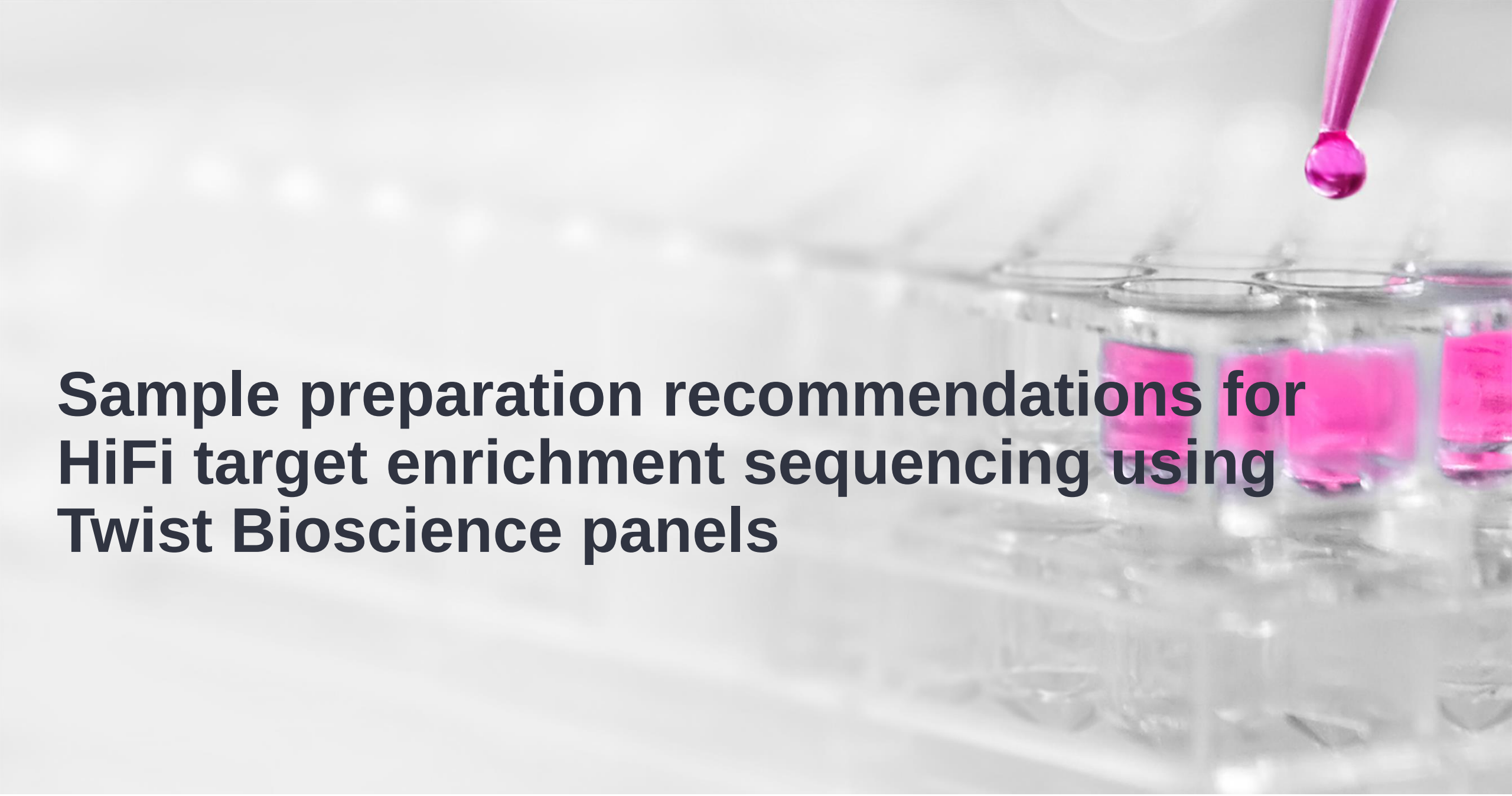
**Full typing result**

| *Allele 1* | *Allele 2* | *CWD 1* | *CWD 2* | *Review status* |
|---|---|---|---|---|
| **B*14:02:01:01** | **B*35:08:01:01** | C | C | Not reviewed |
| Core+ mm: 0<br>Exon+ mm: 0<br>Intron mm: 0 | Core+ mm: 0<br>Exon+ mm: 0<br>Intron mm: 0 | | | |

**Visualization**

Continuous red line marking indicates the HLA data are **fully phased**





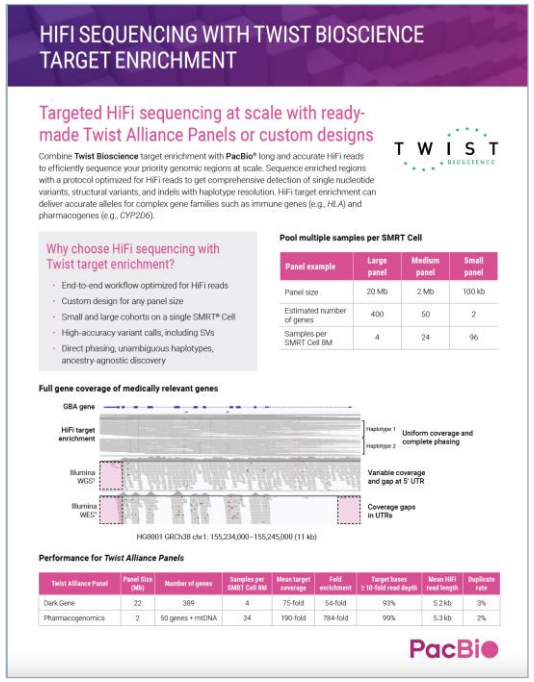
# Sample preparation recommendations for HiFi target enrichment sequencing using Twist Bioscience panels

# Multiplexed HiFi target enrichment sequencing: Technical resources

**Application-specific educational literature**

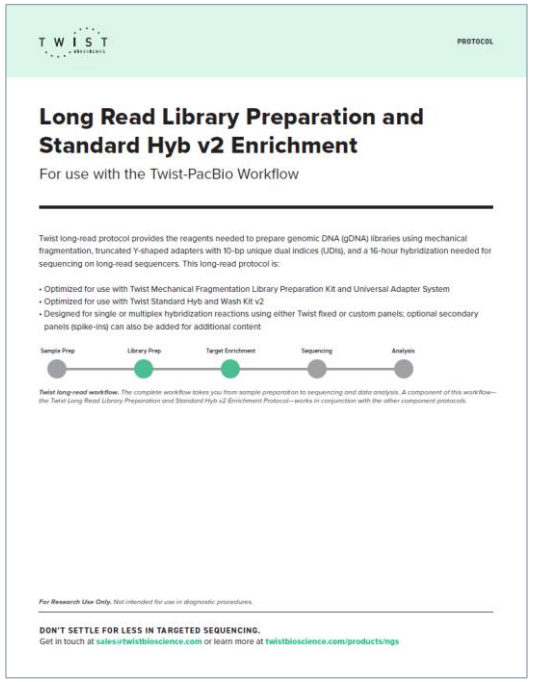
**Application-specific protocol documentation**

**Application-specific technical overviews**



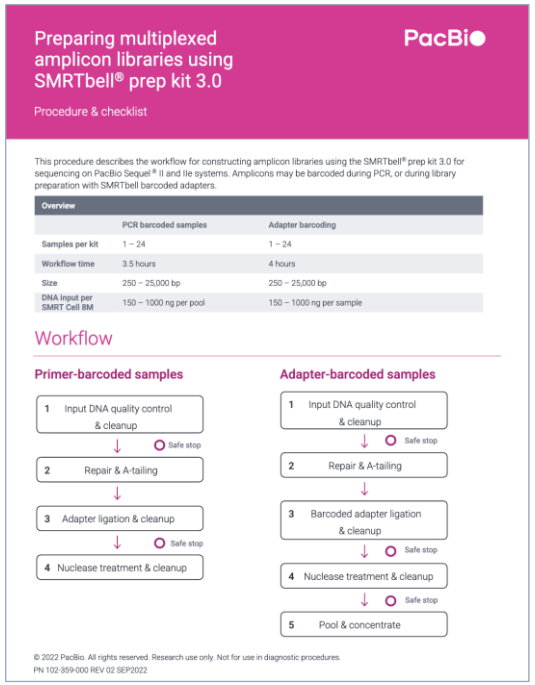

**PacBio literature website ( Link )**

Application-specific brochures, informational guides and other product literature containing best practices recommendations for sample preparation and data analysis workflows.



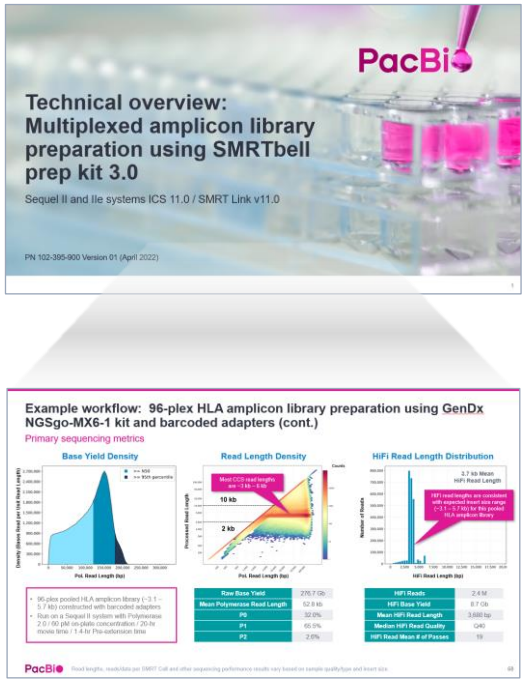

**Twist long read library preparation and standard hyb v2 enrichment** (DOC-001320) [Twist Bioscience]

Technical documentation containing Twist target capture enrichment protocol details.




**Procedure & checklist – Preparing multiplexed amplicon libraries using SMRTbell Prep Kit 3.0** (102-359-000) [PacBio]

Technical documentation containing PacBio SMRTbell library construction details.



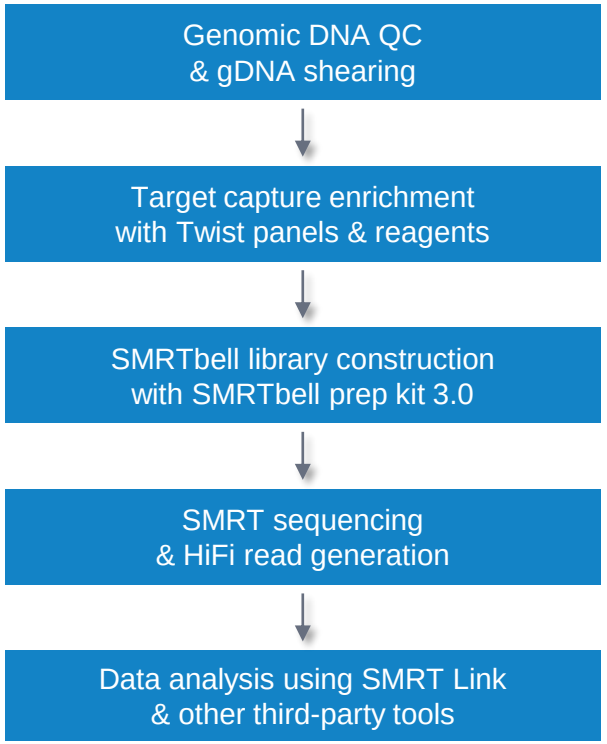

**Technical overview: Multiplexed amplicon library preparation using SMRTbell prep kit 3.0** (102-395-900)

Technical overview presentations describe the end-to-end sample preparation, sequencing setup and data analysis workflow for specific applications. Example sequencing performance data are also summarized.

https://www.pacb.com/connect/datasets/#targeted-datasets

# SMRTbell library preparation workflow overview for HiFi target enrichment samples generated with Twist Bioscience panels

## Follow Twist's long-read target capture protocol (DOC-001320) to generate multiplexed, barcoded amplicon samples suitable for SMRTbell library construction and sequencing on Sequel II and IIe systems




**Workflow for sequencing target enrichment amplicon samples generated with Twist panels.**

- For guidance on performing target capture enrichment of genomic DNA samples, refer to Twist Long read library preparation and standard hyb v2 enrichment (DOC-001320):

| Twist panel type | Large panel | Medium panel | Small panel |
|---|---|---|---|
| Panel size | 20 Mb | 2 Mb | 100 kb |
| No. of genes | 400 | 50 | 2 |
| Samples per SMRT Cell 8M | 4 | 24 | 96 |

  - Start with **500–1000 ng** of high-quality input genomic DNA per sample (50% ≥ 30 kb & 90% ≥ 10 kb)
  - Shear gDNA to target fragment size = **7–10** kb using a Megaruptor 3 system (Diagenode) or g-TUBE (Covaris)
  - Use **200–1000 ng** of sheared gDNA as input for End repair and dA-tailing step
  - **Pool up to 8 samples per hybridization capture reaction** (e.g., for a 24-plex drug panel experiment design, perform 3 parallel hybridization capture reactions containing 8 samples each)
- Barcoded target enrichment amplicon samples may be pooled for SMRTbell library construction and sequencing on a single SMRT Cell 8M by following PacBio Procedure & checklist – Multiplexed amplicon library preparation using SMRTbell Prep Kit 3.0 (102-359-000) (follow workflow instructions for processing primer barcoded samples):
- HiFi data analysis can be performed using SMRT Link and other third-party tools available on GitHub
  - SMRT Link: Demultiplex barcodes, Mark PCR duplicates, HiFi mapping
  - GitHub: Picard CollectHsMetrics, DeepVariant, WhatsHap

# Multiplexed SMRTbell library construction workflow using barcoded target enrichment amplicon DNA samples

**Procedure & checklist – Preparing multiplexed amplicon libraries using SMRTbell Prep Kit 3.0 (102-359-000)**

**Pool barcoded target enrichment amplicon DNA samples**

**1.3X SMRTbell bead cleanup & DNA concentration QC**

Walk-away time

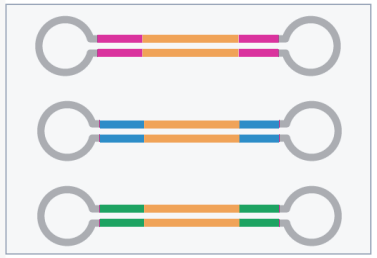
Safe Stop Point

2 hrs (Excl. DNA QC)

35 min — **DNA repair & A-tailing**

Non-barcoded adapter

50 min — **Non-barcoded adapter ligation & 1.3X SMRTbell bead cleanup**



Safe Stop Point

35 min — **Nuclease Treatment & 1.3X SMRTbell bead cleanup**

DNA QC (Sizing & Qubit)

Safe Stop Point

**Sequencing preparation & data collection**

## Pool barcoded amplicon DNA samples generated from Twist target capture workflow (Twist DOC-001320)

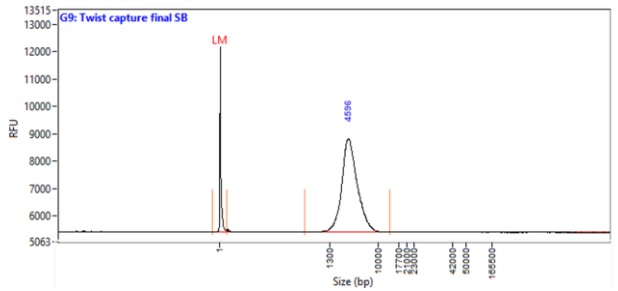
- When amplicons are similar in size, pool an **equal mass** of each sample
- For amplicons <10 kb, recommended total amount of barcoded target enrichment DNA to pool per SMRT Cell 8M for SMRTbell library construction is **150 – 1000 ng***

## Purify and concentrate pooled amplicon DNA sample

- Purify and concentrate pooled DNA sample using **1.3X SMRTbell cleanup beads**
- Elute pooled DNA sample in 47 µL of **low TE buffer**
- Perform **DNA concentration QC** analysis of purified DNA sample (DNA sizing QC is optional)

## Construct multiplexed amplicon SMRTbell library

- Library construction yield with **SMRTbell prep kit 3.0** (102-182-700) is typically ~30 – 50%
- Elute final multiplexed SMRTbell library in 15 µL of **elution buffer (EB)**
- Perform **DNA concentration QC and DNA sizing QC** analysis of final SMRTbell library




## Anneal / bind / cleanup (ABC) with Sequel II binding kit 3.2

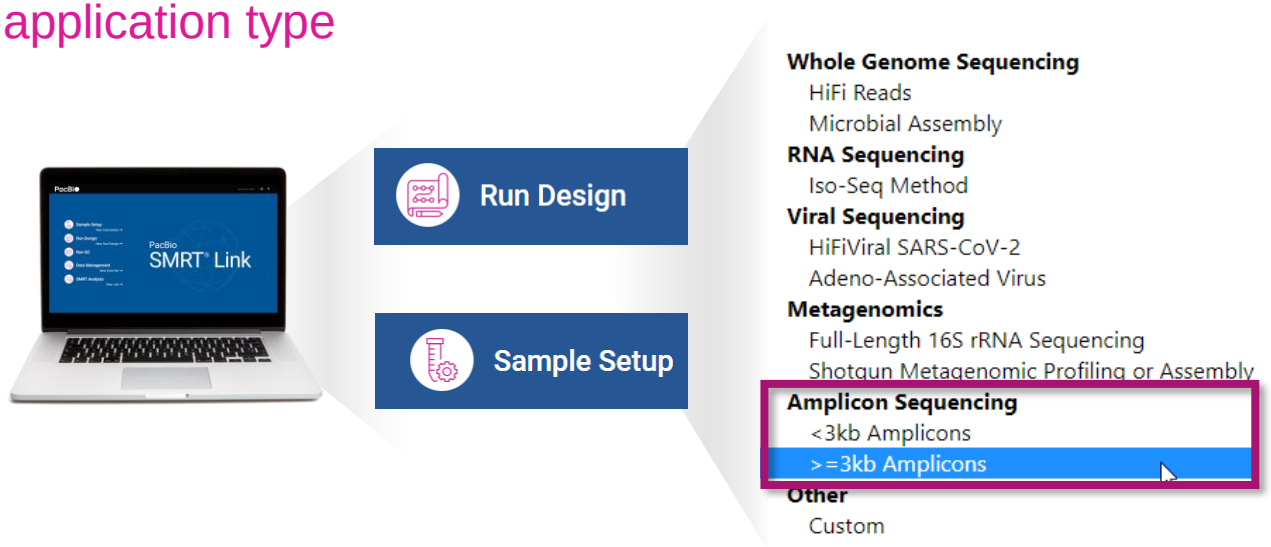
- Use **SMRT Link v11.1 Sample Setup High-Throughput (HT) mode** – Select '≥3 kb Amplicons' application type and perform ABC with **Binding kit 3.2.** (Use Binding kit 2.2 if using SMRT Link v10.2.)
- **≥85 pM** OPLC and **24-hour** movie collection time are recommended

PacBio

* Lower total (pooled) input DNA amounts are possible but may result in lower sequencing yield if there is not enough SMRTbell library available for optimal SMRT cell loading. Refer to Procedure & checklist (102-359-000) for the most up-to-date guidance on recommended DNA input amounts for constructing multiplexed amplicon SMRTbell libraries.

# Sample Setup & Run Design recommendations for HiFi target enrichment libraries

## In SMRT Link Sample Setup & Run Design, select ‘Amplicon Sequencing’ and choose ‘≥3 kb Amplicons’ for application type





**Binding kit 3.2 & cleanup beads** (102-333-300) is recommended for preparing **Twist target enrichment** samples for SMRT sequencing.

- After specifying your application type, **SMRT Link** auto-fills selected Sample Setup and Run Design parameter fields with default recommended values*

| Amplicon library type | Recommended binding kit |
|---|---|
| Twist target enrichment gDNA library | Sequel II binding kit 3.2 |

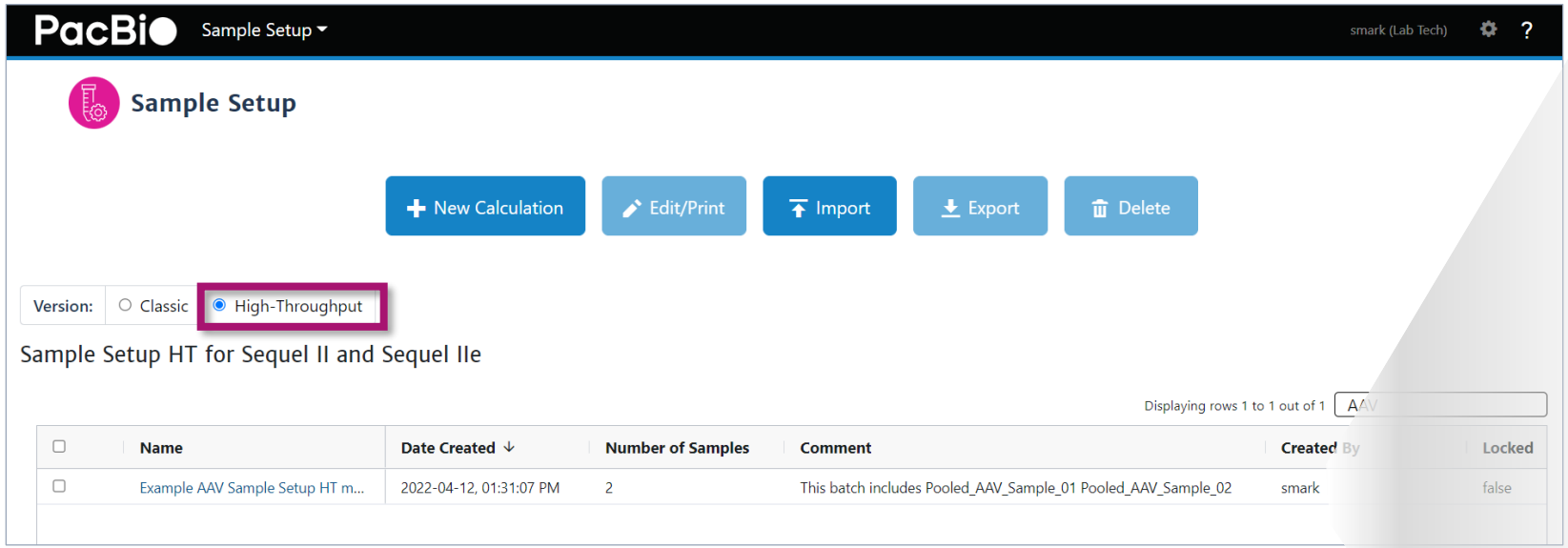
**Sequel II binding kit 3.2 & cleanup beads** (102-333-300) includes the following components:

- Sequencing primer 3.2
- Sequel II polymerase 2.2
- SMRTbell cleanup beads for complex cleanup
- DNA internal control 3.2 (defined 11 kb template bound to Polymerase 2.2)
- Supports ≥24 binding reactions, and up to 4 SMRT Cells 8M per binding reaction (96 cells total), depending on use case, sample size and concentration

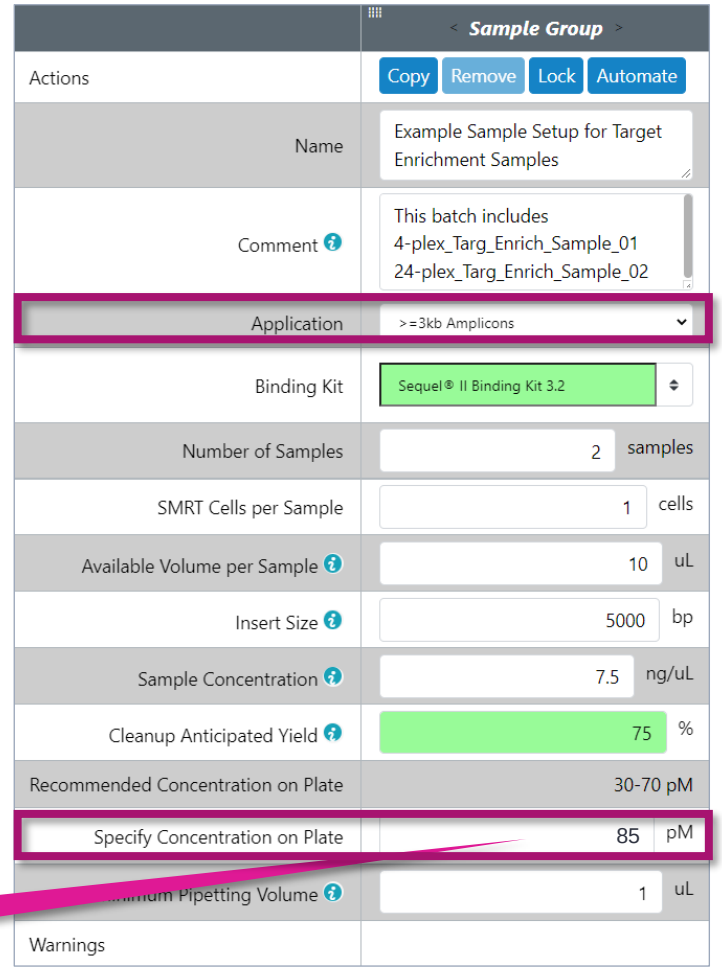
* See **Quick reference card – loading and pre-extension recommendations for Sequel II and IIe systems** (101-769-100) for any updates to sample setup and run design recommendations for specific applications.

# Sample Setup guidance for HiFi target enrichment libraries

## Use SMRT Link Sample Setup High-Throughput (HT) mode and follow instructions to perform ABC (anneal primer / bind polymerase / clean up complex) by selecting '≥3 kb Amplicons' for application type




- Sample Setup High-Throughput mode provides a simplified, streamlined workflow to efficiently process either one sample or multiple samples with similar library properties (such as mean insert size and DNA concentration) in parallel
- You can also export the calculated values to a CSV file for laboratory automation




**Example Sample Setup HT mode worksheet for a batch comprised of two target enrichment amplicon samples.**

**Note:** We recommend starting with an on-plate loading concentration (OPLC) of **≥85 pM** for ≥3 kb target enrichment amplicon samples and adjusting higher or lower if needed to achieve optimal *P1* loading.

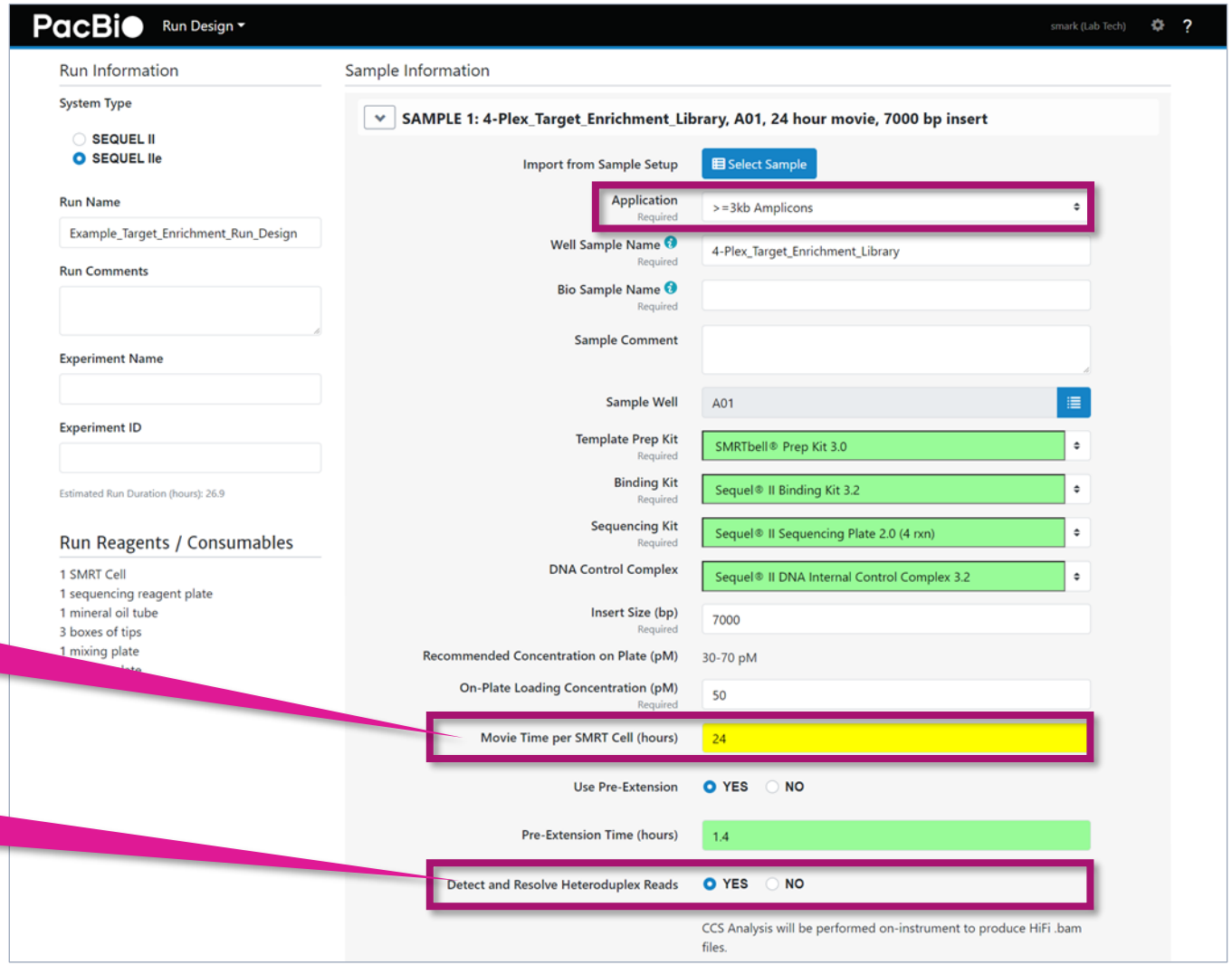
# Run Design guidance for HiFi target enrichment libraries

## Follow SMRT Link Run Design instructions to set up a sequencing run by selecting '≥3 kb Amplicons' for application type

- Select ≥3 kb Amplicons from the Application field drop-down menu in SMRT Link Run Design
- The following fields are auto-populated and high-lighted in green:
  - Template Prep Kit
  - Binding Kit
  - Sequencing Kit
  - DNA Control Complex
  - Movie Time Per SMRT Cell
  - Pre-Extension Time

For HiFi target enrichment samples, we recommend using a **24-hour** movie collection time

By default, heteroduplex detection and splitting feature is automatically set to **YES** for ≥3 kb Amplicon samples.




**Example sample information entered into Run Design for sequencing a 4-plex pooled HiFi target enrichment amplicon sample.**

# Run Design guidance for HiFi target enrichment libraries (cont.)

We recommend setting up your SMRT Link Run Design to specify Sample Is Barcoded = NO and manually performing barcode demultiplexing in SMRT Link using an appropriate barcode FASTA file

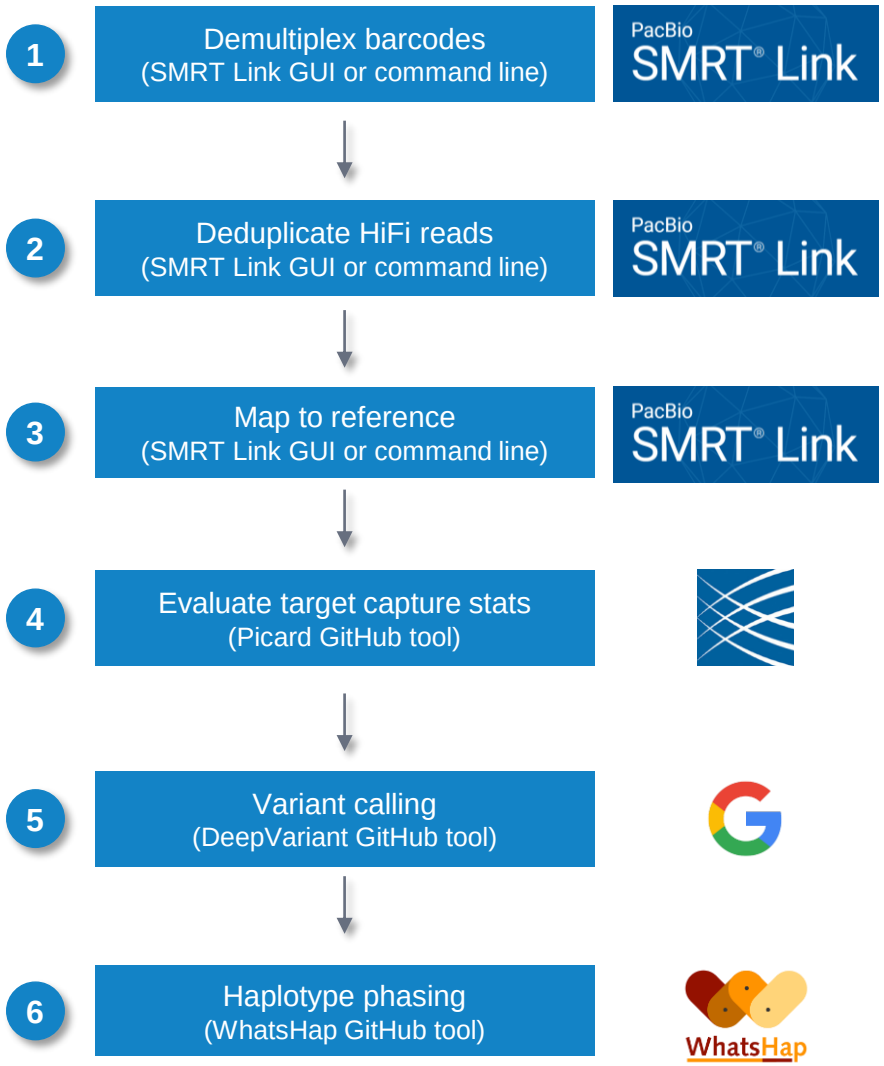

- In SMRT Link Run Design, specify Sample Is Barcode = NO
- Manually perform barcode demultiplexing of target enrichment data using SMRT Link Demultiplex Barcodes data utility (via GUI or command line)
  - Use the Twist TruSeq barcode FASTA file that contains the barcode and Twist universal adapter sequences (available on our Multiplexing website*).



* Refer to our Multiplexing website for application-specific recommendations for barcodes, procedures, ordering sheets and FASTA files required for sample demultiplexing.

# HiFi target enrichment data analysis workflow recommendations

## Use SMRT Link and other third-party tools for performing data analysis QC




Separate HiFi reads by sample barcode using SMRT Link Demultiplex Barcodes data utility

- Command line tool: `lima` (https://github.com/PacificBiosciences/barcoding)

Deduplicate HiFi reads using SMRT Link Mark PCR Duplicates data utility

- Command line tool: `pbmarkdup` (https://github.com/PacificBiosciences/pbmarkdup)

Map HiFi reads to reference using SMRT Link HiFi Mapping analysis application

- Command line tool: `pbmm2` (https://github.com/PacificBiosciences/pbmm2)

Evaluate on-target capture statistics using Picard GitHub tool

- Command line tool: `picard` (https://snakemakewrappers.readthedocs.io/en/stable/wrappers/picard/collecthsmetrics.html)

Perform variant calling using DeepVariant GitHub tool

- Command line tool: `deepvariant` (https://github.com/google/deepvariant)

Perform haplotype phasing using WhatsHap GitHub tool

- Command line tool: `whatshap` (https://whatshap.readthedocs.io/en/latest/)



For advanced users, a command line workflow using developmental tools is available through GitHub – Please contact PacBio Technical Support to request access.

# Example: 4-Plex Twist Alliance Dark Genes Panel

## Data analysis and QC results

- Comprehensive 22 Mb panel: Full gene coverage for 389 **challenging medically-relevant genes**[1]
- Uncover genes in "**NGS dead zones**" that are difficult to sequence or map with short reads[2,3]

### Sequencing metrics

| | |
|---|---|
| **OPLC** | 85 pM |
| **P1 %** | 79.7 % |
| **HiFi reads** | 3.76 M |
| **HiFi base yield per SMRT Cell 8M** | 19.53 Gb |
| **HiFi read length (mean)** | 5193 bp |
| **HiFi read quality (median)** | Q40 |

### Analysis metrics

| | |
|---|---|
| **Panel Size** | 22 Mb |
| **Number of genes** | 389 |
| **Sample per SMRT Cell 8M** | 4 |
| **Mean reads per sample** | 893,459 |
| **Mean target coverage** | 75-fold |
| **Target bases at ≥ 10-fold** | 93% |
| **Fold enrichment** | 54-fold |
| **PCR duplicate rate** | 2.78% |


[1] Ji *et al.* Characterizing the genetic polymorphisms in 370 challenging medically relevant genes using long-read sequencing data from 41 human individuals among 19 global populations. bioRXiv https://doi.org/10.1101/2022.08.03.502734

[2] Mandelker *et al.* Navigating highly homologous genes in a molecular diagnostic setting: a resource for clinical next-generation sequencing. Genetics in Medicine 2016.

[3] Wenger *et al.* Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. Nature Biotech (2019)

# Example: 24-Plex Twist Alliance PGx Panel

## Data analysis and QC results

- Comprehensive 50 gene 2 Mb panel
- 39 full-length genes enable phasing
- Includes all 20 genes with CPIC guidelines[1]

### Sequencing metrics

| | |
|---|---|
| **OPLC** | 85 pM |
| **P1 %** | 76.3 % |
| **HiFi reads** | 3.76 M |
| **HiFi base yield per SMRT Cell 8M** | 20.11 Gb |
| **HiFi read length (mean)** | 5348 bp |
| **HiFi read quality (median)** | Q42 |

### Analysis metrics

| | |
|---|---|
| **Panel Size** | 2 Mb |
| **Number of genes** | 50 |
| **Sample per SMRT Cell 8M** | 24 |
| **Mean reads per sample** | 149,749 |
| **Mean target coverage** | 190-fold |
| **Target bases at ≥ 20-fold** | 96% |
| **Fold enrichment** | 784-fold |
| **PCR duplicate rate** | 1.84% |

[1] https://cpicpgx.org/guidelines/

# Technical documentation & applications support resources

# Technical resources for multiplexed amplicon library preparation, sequencing & data analysis

## Sample & library preparation literature

- Application brief – HiFi target enrichment – Best practices (102-326-515)
- Application brief – Targeted sequencing for amplicons – Best practices (102-193-603)
- Application note – HiFi amplicon sequencing for pharmacogenetics: CYP2D6 (102-326-548)
- Application note – Targeted HiFi sequencing for congenital adrenal hyperplasia (102-326-547)
- Application note – Targeted HiFi sequencing for thalassemia (102-326-551)
- Overview – Sequel systems application options and sequencing recommendations (101-851-300)
- Procedure & checklist – Amplification of bacterial full-length 16S rRNA gene with barcoded primers (101-599-700)
- Procedure & checklist – Preparing multiplexed amplicon libraries using PacBio barcoded M13 primers and SMRTbell prep kit 3.0 (101-921-300)
- Procedure & checklist – Preparing multiplexed amplicon libraries using SMRTbell prep kit 3.0 (102-359-000)
- Technical note – Preparing DNA for PacBio HiFi sequencing – Extraction and quality control (102-193-651)
- Technical note – Sample preparation for PacBio HiFi sequencing from human whole blood (102-326-500)
- Technical overview – Multiplexed amplicon library preparation using SMRTbell prep kit 3.0 (102-395-900)
- Twist protocol – Long Read Library Preparation and Standard Hyb v2 Enrichment (DOC-001320)

## Data analysis resources

- SMRT Link user guide (v11.1) (102-413-100)
- SMRT Tools reference guide (v11.1) (102-413-200)

# Technical resources for multiplexed amplicon library preparation, sequencing & data analysis (cont.)

## Posters

- CPIC poster (2022): Enablement of long-read targeted pharmacogenomic panels using Twist hybrid capture and PacBio HiFi sequencing [ Link ]
- ESHG poster (2022): Enablement of long-read targeted sequencing panels using Twist hybrid capture and PacBio HiFi sequencing [ Link ]
- ASHG poster (2021): Long-read amplicon sequencing of the polymorphic CYP2D6 locus [ Link ]
- ASHG poster (2021): Resolving Complex Pathogenic Alleles using HiFi Long-Range Amplicon Data and a New Clustering Algorithm [ Link ]

## Publications

- Van der Lee, Maaike et al. (2022) Design and performance of a long-read sequencing panel for pharmacogenomics. BioRxiv preprint. [ Link ]
- Luo, Shiqiang et al. (2022) Detection of four rare thalassemia variants using Single-molecule real time sequencing. Front. Genet. [ Link ]
- Twesigomwe, David et al. (2022) Characterization of CYP2D6 Pharmacogenetic variation in sub-Saharan African populations. Clin Pharmacol Ther. [ Link ]
- Scott, Erick R et al. (2022) Long-read HiFi sequencing of*NUDT15*: Phased full-gene haplotyping and pharmacogenomic allele discovery. Human Mutation. 43: 1557-1566 [ Link ]
- Te Paske, Iris B.A.W. et al (2022) Non-coding aberrations in mismatch repair genes underlie a substantial part of the missing heritability in Lynch syndrome. Gastroenterology. [ Link ]
- Rodriguez Oscar L et al. (2022) Targeted long-read sequencing facilitates phased diploid assembly and genotyping of the human T cell receptor alpha, delta and beta loci. bioRXiv preprint [ Link ]

# Technical resources for multiplexed amplicon library preparation, sequencing & data analysis (cont.)

## Webinars

- PacBio ASHG webinar (2021): Using HiFi reads for improved and accurate haplotyping and phasing of pharmacogenomic alleles [ Link ]
- PacBio webinar (2021): Identifying key players in host-microbiome interactions with high resolution 16S sequencing [ Link ]
- PacBio application tutorial (2020): Introduction to targeted sequencing with HiFi reads [ Link ]

## Example PacBio data sets

| Targeted sequencing application | Dataset* | Data type | PacBio system |
|---|---|---|---|
| PCR amplicon sequencing | CYP2D6 amplicon for PGx reference samples | HiFi Reads | Sequel IIe System |
| HiFi target enrichment with Twist probes | Twist Alliance Long Read PGx Panel for 10 reference samples | HiFi Reads | Sequel IIe System |
| HiFi target enrichment with Twist probes | Twist Alliance Dark Gene Panel for NA12878 | HiFi Reads | Sequel IIe System |
| Full-length 16S sequencing | 20 strain mock microbial community – ATCC MSA-1003 – 16S | HiFi Reads | Sequel II System |
| HLA Sequencing | Analysis of HLA Amplicons (HLA-A, -B, -C, -DRB1, -DQB1 and –DPB1) Generated Using GenDX NGSgo-MX6-1 Kit | HiFi Reads | Sequel II System |

Visit our Datasets website for a complete list of example SMRT sequencing datasets for specific applications.

www.pacb.com