Technical overview: Adeno-associated virus (AAV) library preparation using SMRTbell prep kit 3.0

Sequel II and IIe systems ICS v11.0 / SMRT Link v11.0

PN 102-390-400 Version 01 (April 2022)

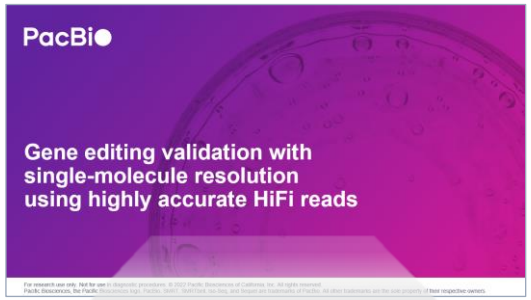
Adeno-associated virus (AAV) library preparation using SMRTbell prep kit 3.0

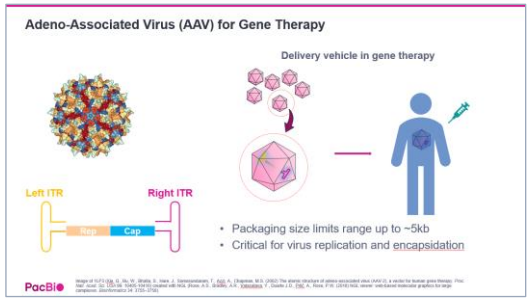
Technical overview

1. PacBio sequencing for quality control of gene therapy methods
2. AAV library sample preparation workflow overview
3. AAV library sample preparation workflow details
4. AAV library sequencing preparation workflow overview
5. AAV data analysis workflow overview
6. AAV library example performance data
7. Technical documentation & applications support resources

AAV Sequencing: How to get started

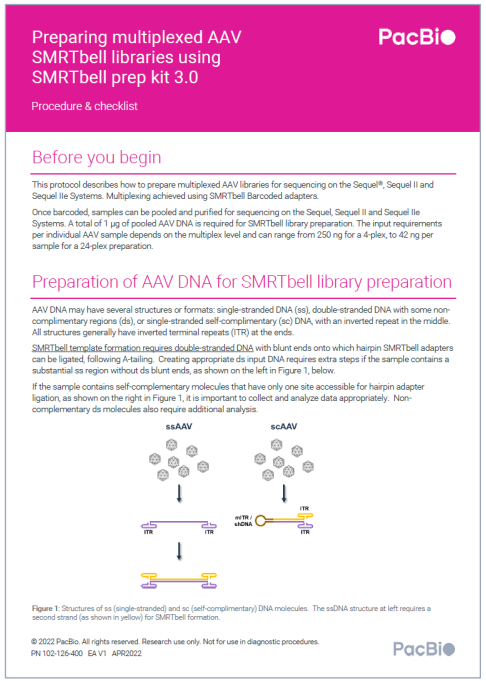
Application-specific educational literature → **Application-specific Procedure & checklist** → **Application-specific technical overviews** → **Library construction, sequencing & analysis**





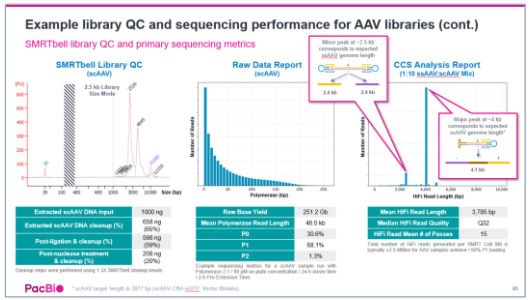
PacBio literature website (Link)

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



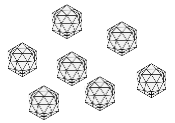
Procedure & Checklist – Preparing multiplexed AAV SMRTbell libraries using SMRTbell prep kit 3.0 (102-126-400)

Technical documentation containing sample library construction and sequencing preparation protocol details



Technical Overview: AAV Library preparation using SMRTbell prep kit 3.0 (102-390-400)

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.



AAV DNA extraction & QC

≥250 ng per AAV sample for a 4-plex; or ≥42 ng per AAV sample for a 24-plex

Library construction (SMRTbell prep kit 3.0)

Multiplex AAV samples using SMRTbell barcoded adapter plate 3.0

Sequencing (Sequel II and IIe systems)

ABC* with Sequel II Binding Kit 3.1

24-h movie collection time

Data analysis (GitHub)

Map HiFi reads to AAV reference and visualize alignment results

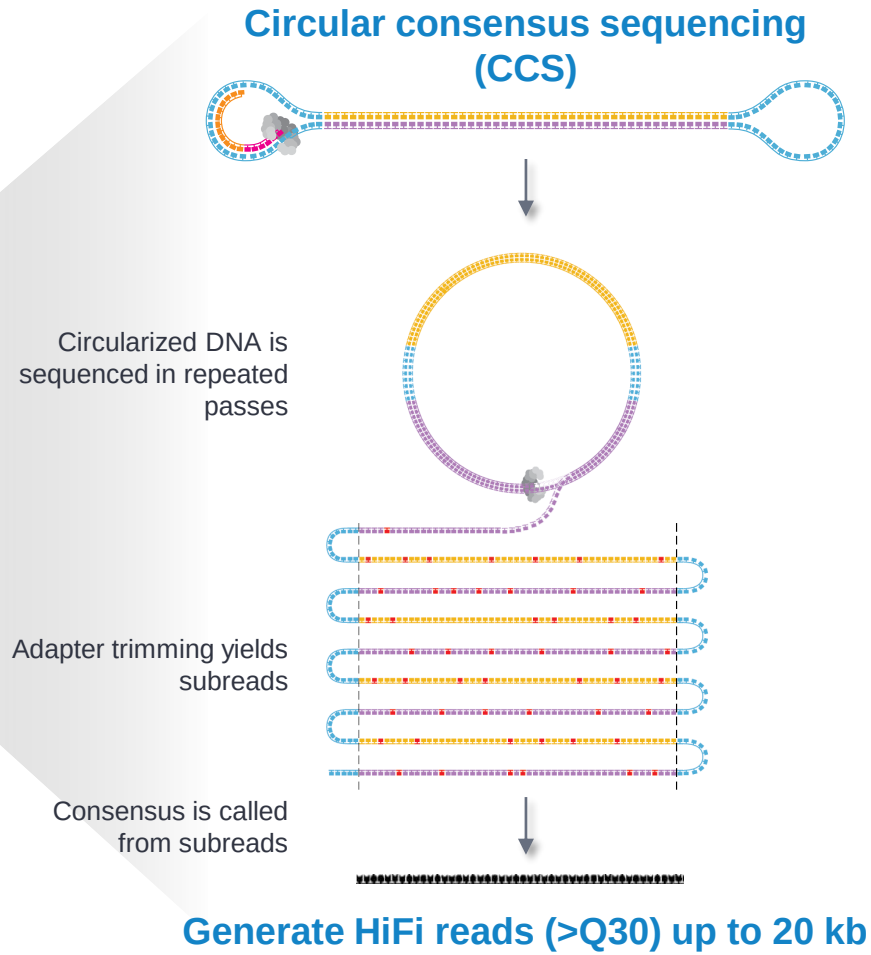
PacBio

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

PacBio sequencing for quality control of gene therapy methods

PacBio sequencing for quality control of gene therapy methods

PacBio is the only sequencing technology able to support a broad range of gene therapy use cases



Generate HiFi reads (>Q30) up to 20 kb

High accuracy and long read lengths enables a broad range of gene therapy methods

- **Discovery & Engineering**
- **Vector QC**
- **Editing Efficiency**
- **Safety**

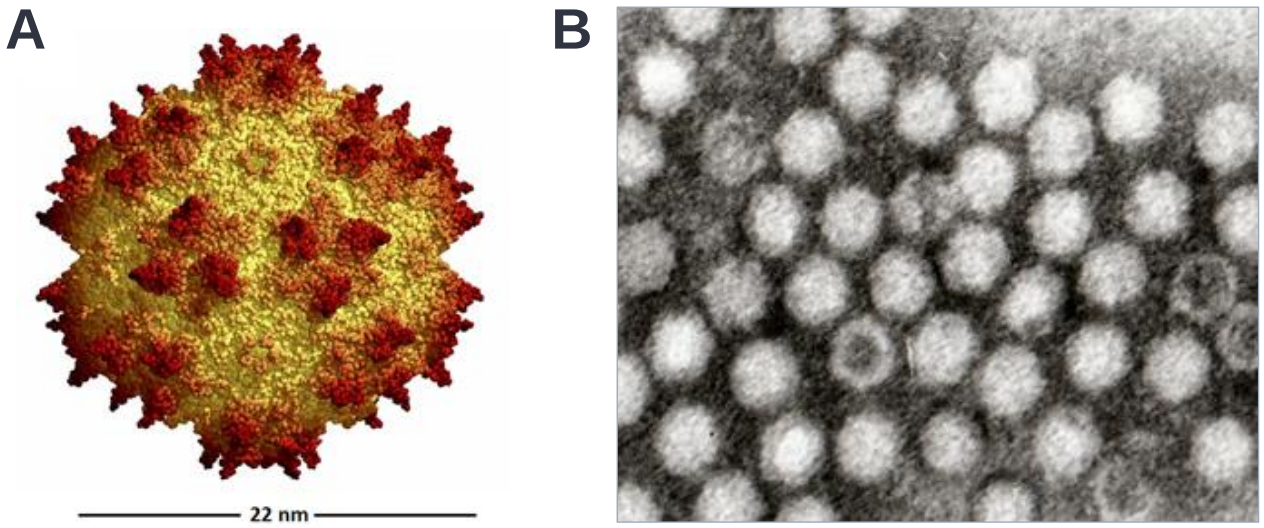
Adeno-associated virus (AAV) for gene therapy

Recombinant adeno-associated viral (rAAV) vectors are promising tools for gene therapy

AAV structure is comprised of a non-enveloped capsid protein shell surrounding and protecting a small, single-stranded DNA genome of ~4.8 kb

Why Use AAV as a Vector

- Ability to generate recombinant AAV (rAAV)
- Form episomal concatemers in the cell nucleus
- Reduced host genome integration
- Proven in various therapeutic applications
- Low immunogenicity
- Strong vector persistence
- Long-term transferred gene (transgene) expression
- Non-toxic



Adeno-Associated Virus (AAV) structure. **A.** Cartoon illustration showing simulated AAV size and 3-D structure. **B.** Electron microscope image of purified AAV vector particles.

Images from "Adeno-Associated Virus (AAV) Structure and Genome map." Introduction of AAV. http://signagen.com/Introduction-to-AAV. Accessed April.19.2022.

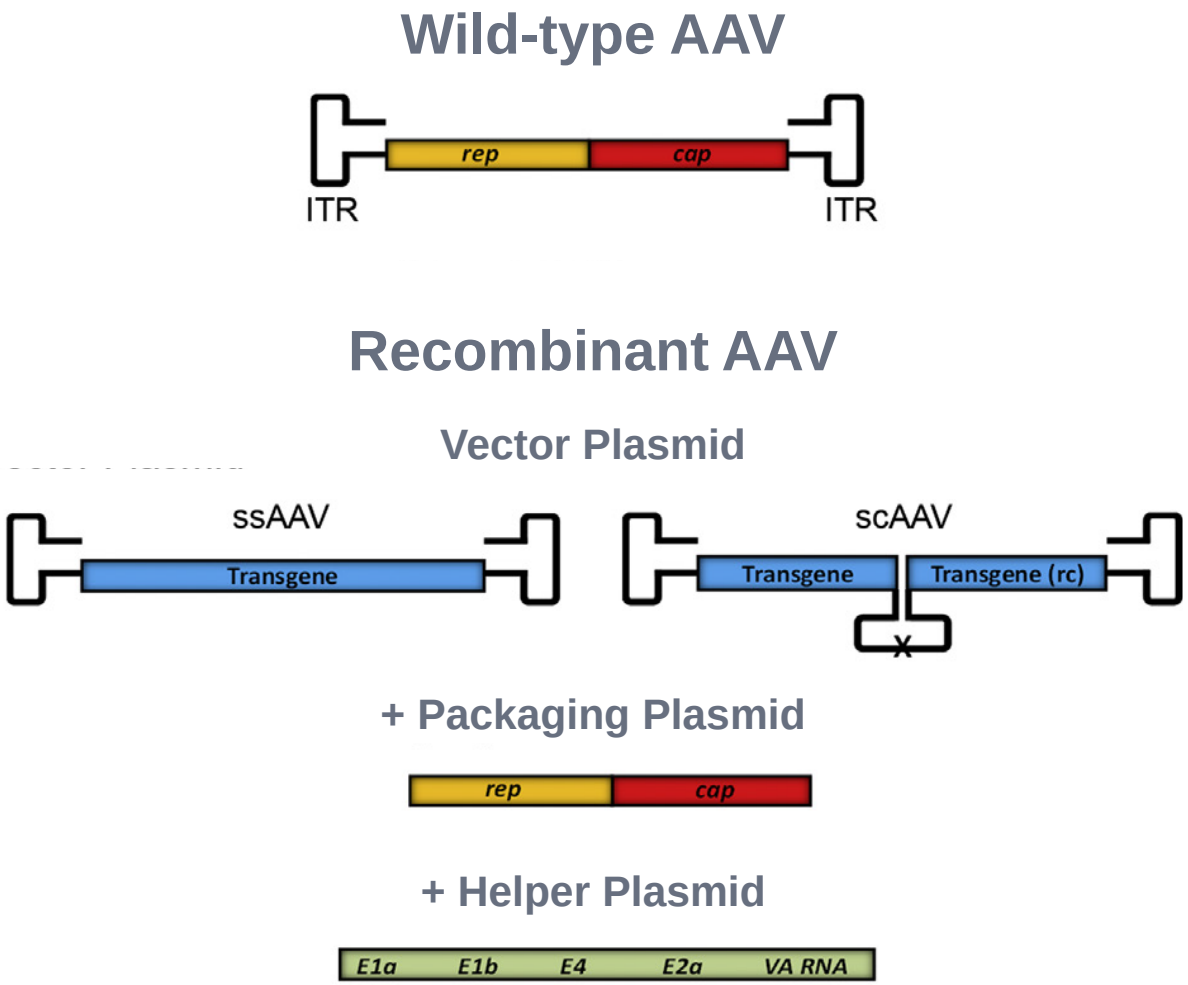
Recombinant AAV vector genome design

Single-stranded AAV (ssAAV) and self-complementary AAV (scAAV)

- Wild-type AAV genome consists of the viral *rep* and *cap* genes between two **inverted terminal repeats (ITRs)**
- Recombinant AAV vector is produced by co-transfection of
    1. **Vector plasmid** containing an ITR-flanked transgene cassette;
    2. **Packaging plasmid** that encodes the *rep* and *cap* genes of a specific AAV serotype; and
    3. **Helper plasmid** that supplies the essential adenovirus helper genes

The vector plasmid may encode for:

- **Single-stranded DNA (ssAAV)** containing ~4.5 kb of novel transgene sequence; or
- **Self-complementary DNA (scAAV)** containing up to ~2.2 kb of novel transgene sequence in duplex form



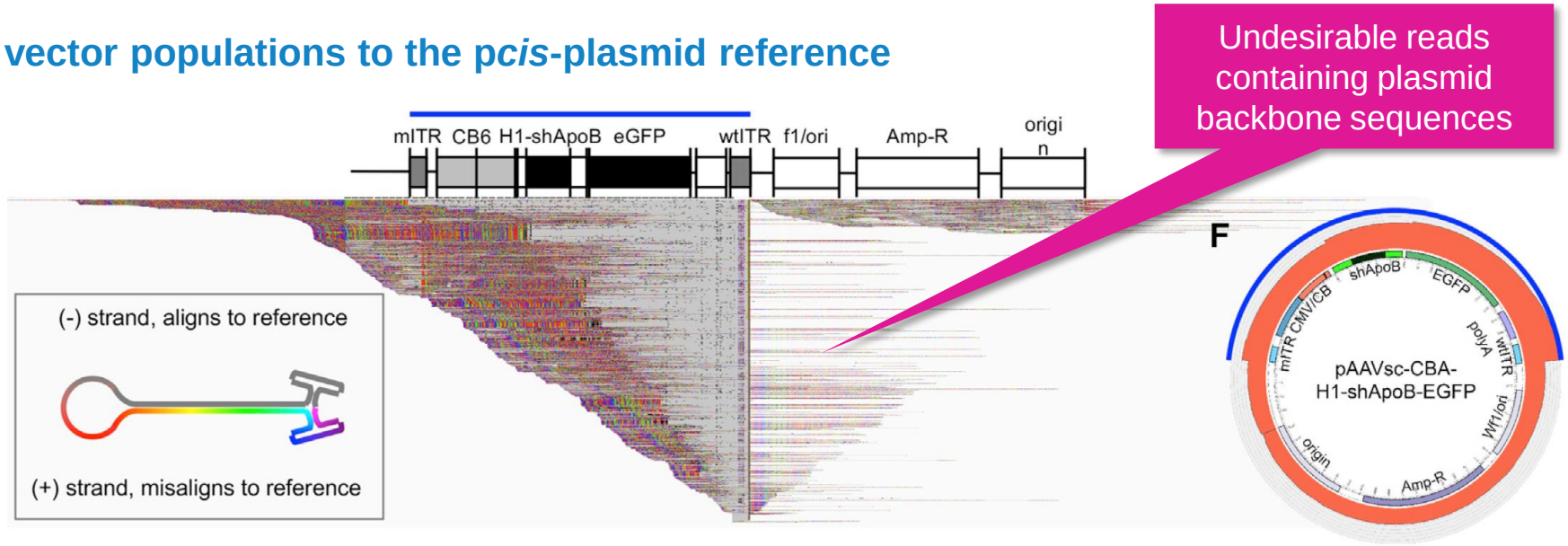
Kantor et al. (2014) Methods for Gene Transfer to the Central Nervous System. Adv Genet. 87:125-97

Recombinant AAV Vector QC Example

PacBio HiFi reads allow for complete sequencing of viral vector particles. This allows for unambiguous detection of undesired plasmid or host gDNA that has been encapsidated

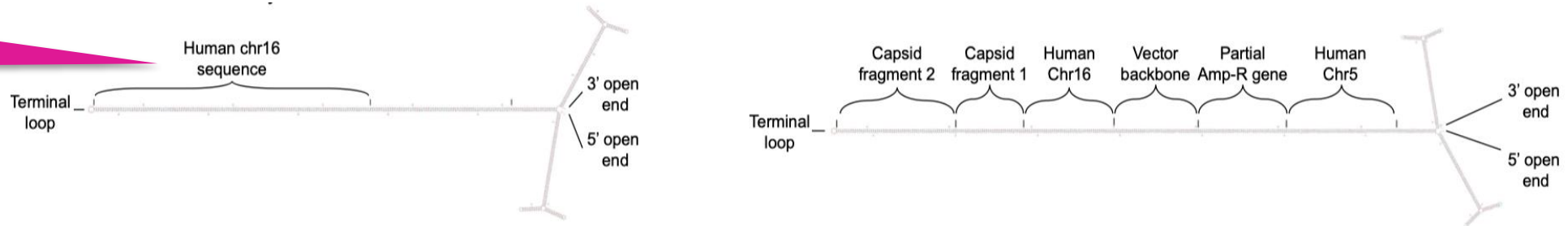
Alignments of heterogeneous vector populations to the p*cis*-plasmid reference

Vector QC is central to understanding that the vectors are packaging the **correct** DNA



Characterization of chimeric reads that map to the human genome

Undesirable reads containing human host sequences



Tai et al. (2018) Adeno-associated Virus Genome Population Sequencing Achieves Full Vector Genome Resolution and Reveals Human-Vector Chimeras. Molecular Therapy - Methods & Clinical Development. 9:130-141

AAV library sample preparation workflow overview

AAV sample preparation procedure description

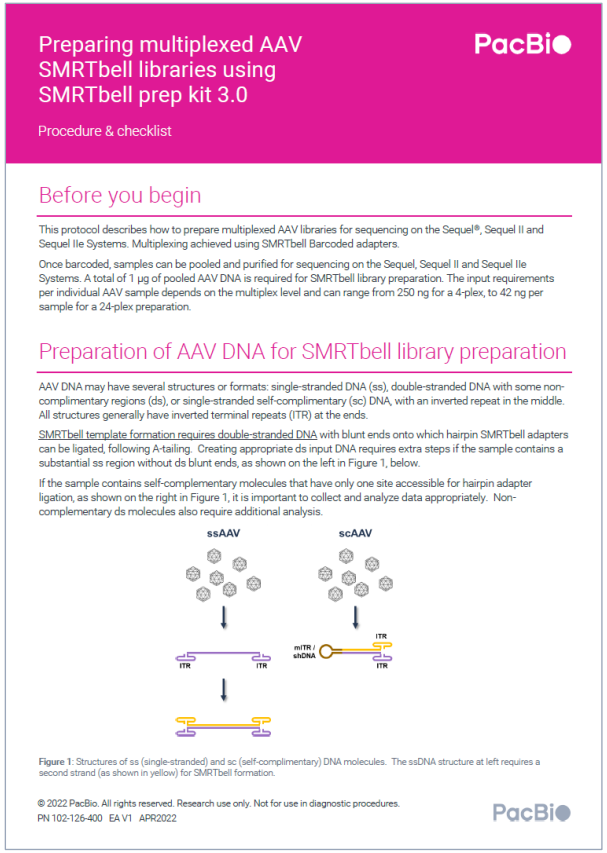
Procedure & checklist – Preparing multiplexed adeno-associated virus (AAV) libraries using SMRTbell prep kit 3.0 (102-126-400) describes a library preparation procedure for sequencing multiplexed AAV samples on the Sequel II and IIe Systems

SMRTbell Prep Kit 3.0
(102-182-700)

SMRTbell Barcoded Adapter Plate 3.0
(102-009-200)

Procedure & checklist highlights

- Compatible with single-stranded AAV (ssAAV) and self-complementary AAV (scAAV)
- AAV DNA sample extraction is performed using third-party methods
- Multiplexing is performed using SMRTbell barcoded adapter plate 3.0 (102-009-200)
- A total of 1 µg of pooled AAV DNA is required for SMRTbell library preparation. The input DNA requirement per individual AAV sample depends on the multiplex level and can range from 250 ng per sample for a 4-plex, to 42 ng per sample for a 24-plex preparation.



Preparing multiplexed AAV SMRTbell libraries using SMRTbell prep kit 3.0

PacBio

Procedure & checklist

Before you begin

This protocol describes how to prepare multiplexed AAV libraries for sequencing on the Sequel®, Sequel II and Sequel IIe Systems. Multiplexing achieved using SMRTbell Barcoded adapters.

Once barcoded, samples can be pooled and purified for sequencing on the Sequel, Sequel II and Sequel IIe Systems. A total of 1 µg of pooled AAV DNA is required for SMRTbell library preparation. The input requirements per individual AAV sample depends on the multiplex level and can range from 250 ng for a 4-plex, to 42 ng per sample for a 24-plex preparation.

Preparation of AAV DNA for SMRTbell library preparation

AAV DNA may have several structures or formats: single-stranded DNA (ss), double-stranded DNA with some non-complimentary regions (ds), or single-stranded self-complimentary (sc) DNA, with an inverted repeat in the middle. All structures generally have inverted terminal repeats (ITR) at the ends.

SMRTbell template formation requires double-stranded DNA with blunt ends onto which hairpin SMRTbell adapters can be ligated, following A-tailing. Creating appropriate ds input DNA requires extra steps if the sample contains a substantial ss region without ds blunt ends, as shown on the left in Figure 1, below.

If the sample contains self-complementary molecules that have only one site accessible for hairpin adapter ligation, as shown on the right in Figure 1, it is important to collect and analyze data appropriately. Non-complementary ds molecules also require additional analysis.

Figure 1: Structures of ss (single-stranded) and sc (self-complimentary) DNA molecules. The ssDNA structure at left requires a second strand (as shown in yellow) for SMRTbell formation.

PN 102-126-400 EA V1 APR2022

PacBio

PacBio Documentation (102-126-400)

APPLICATIONS
VIRAL SEQUENCING
AAV Sequencing

AAV sample preparation & sequencing workflow overview

Workflow summary for constructing SMRTbell libraries suitable for sequencing on the Sequel and Sequel II/IIe Systems for multiplexed AAV applications

AAV DNA extraction & QC

- Perform AAV DNA sample extraction using third-party methods (Samples can be comprised of a mixture of ssAAV and scAAV species)
- ≥1 µg of pooled AAV DNA is required per SMRT Cell 8M
- Input DNA requirement per AAV sample can range from 250 ng for a 4-plex, to 42 ng per sample for a 24-plex.

SMRTbell library construction

- *Procedure & checklist – Preparing multiplexed adeno-associated virus (AAV) libraries using SMRTbell prep kit 3.0* (102-126-400)
- Multiplex AAV samples using SMRTbell Barcoded Adapter Plate 3.0 (102-009-200).
- Pool barcoded AAV samples and purify final multiplexed SMRTbell library using SMRTbell cleanup beads

Sequencing

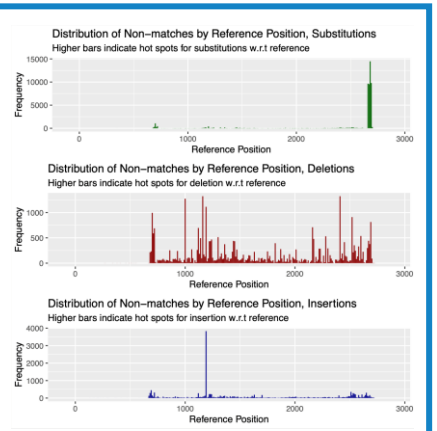
- Follow SMRT Link Sample Setup instructions for primer annealing, polymerase binding, complex cleanup and sample loading

HiFi Read

PacBio HiFi reads achieve 99.9% accuracy

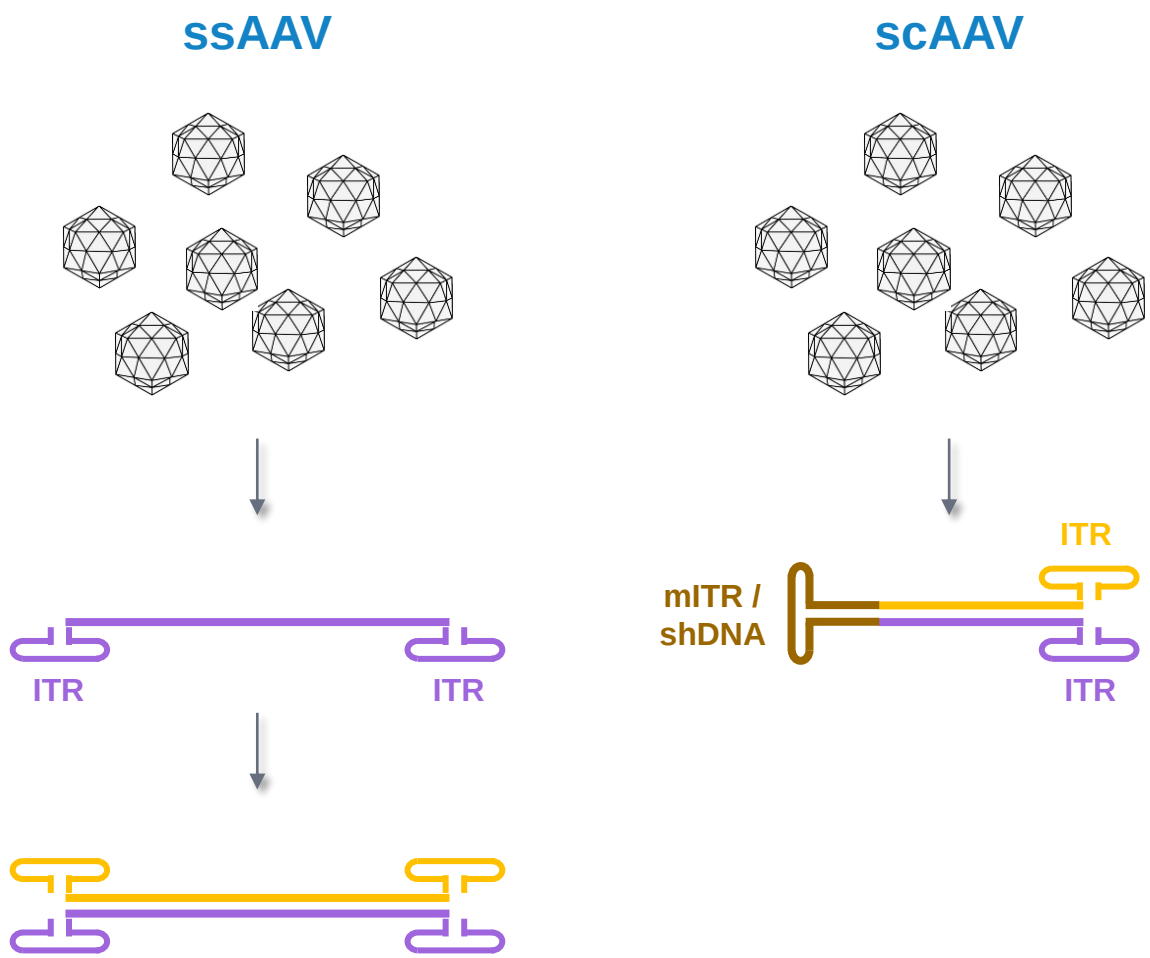
Data analysis

- Perform CCS analysis and de-multiplex barcodes on-instrument (Sequel IIe system only) or in SMRT Link
- Align demultiplexed AAV data against a reference and visualize results using command line tools available in PacBio GitHub



AAV DNA extraction for SMRTbell library preparation

- Refer to Procedure & checklist – Preparing multiplexed adeno-associated virus (AAV) libraries using SMRTbell prep kit 3.0 (102-126-400) for recommendations on third-party AAV DNA extraction methods
- AAV DNA may have several structures or formats:
  - Single-stranded (ss) DNA;
  - Double-stranded (ds) DNA with some non-complementary regions; or
  - Single-stranded self-complementary (sc) DNA, with an inverted repeat in the middle
- All structures generally have inverted terminal repeats (ITR) at the ends (some may have mutated ITRs or other short hairpin sequences present)
- **Note:** SMRTbell template synthesis requires dsDNA with blunt ends onto which hairpin SMRTbell adapters can be ligated
  - Creating appropriate dsDNA input for SMRTbell library construction requires an extra thermal annealing step if the sample contains substantial ss regions without ds blunt ends, as shown on the left in the figure for ssAAV samples



Structures of single-stranded (ss) (left) and self-complementary (sc) (right) AAV DNA molecules. The ssDNA structure at left requires a second strand (as shown in yellow) for SMRTbell library construction.

AAV SMRTbell library construction workflow overview

DNA repair & A-tailing

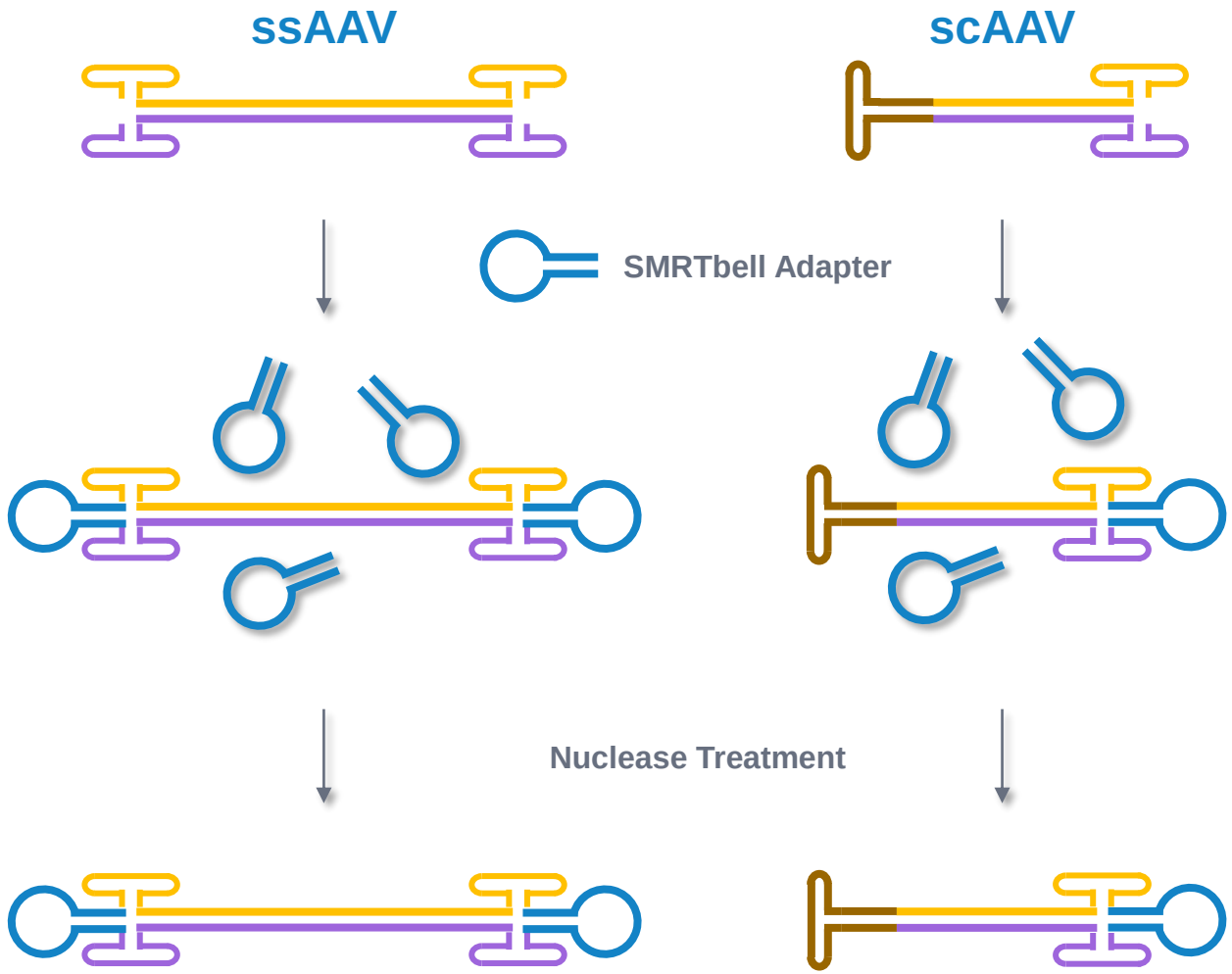
- After extracting AAV DNA, perform DNA damage repair, polish fragment ends and carry out A-tailing reaction using SMRTbell prep kit 3.0 (102-182-700)

Barcoded adapter ligation & cleanup

- Use SMRTbell barcoded adapter plate (102-009-200) to prepare multiplexed AAV samples

Nuclease treatment & cleanup

- Perform nuclease treatment and cleanup to remove incomplete SMRTbell templates/un-ligated adapters and obtain a purified AAV library sample ready for pooling & sequencing



See **Procedure & checklist – Preparing multiplexed adeno-associated virus (AAV) libraries using SMRTbell prep kit 3.0** (102-126-400) for details.

Structure of AAV SMRTbell library template molecules

Comparison of standard SMRTbell library template structure *versus* ssAAV SMRTbell templates and scAAV SMRTbell templates

Standard SMRTbell template

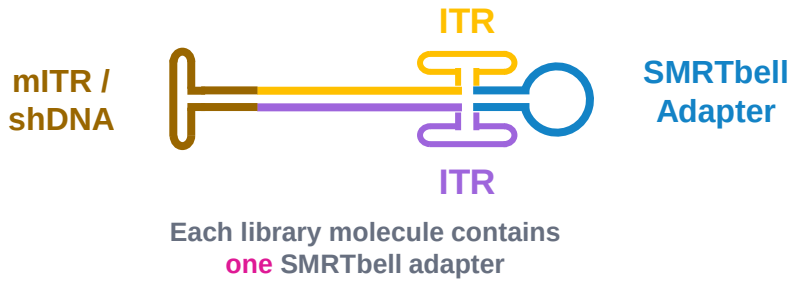
- Linear dsDNA insert molecules are ligated to SMRTbell adapters at each end during library construction

Single-stranded AAV (ssAAV) SMRTbell template

- ssAAV molecules can be converted to dsDNA structures by performing intermolecular annealing or 2nd-strand synthesis
- Forward (+) and reverse (–) sequences are on separate strands, bracketed by T-shaped inverted terminal repeat (ITR) sequences and ligated to SMRTbell adapters at each end during library construction

Self-complementary AAV (scAAV) SMRTbell template

- scAAV molecules feature a linear self-complementary sequence containing a mutant ITR (mITR) (or short hairpin DNA, shDNA) in the middle of the molecule
- The two ITRs at the open end of the scAAV vector genome are ligated to a single SMRTbell adapter



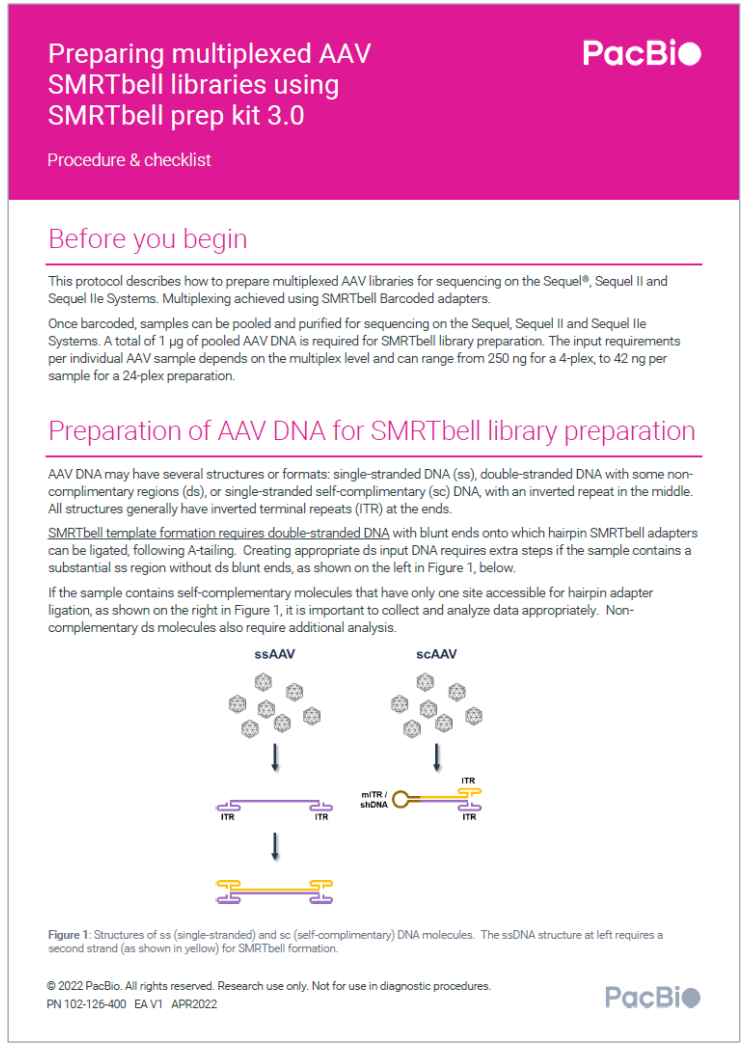
AAV library sample preparation workflow details

Procedure & checklist – Preparing multiplexed AAV SMRTbell libraries using SMRTbell prep kit 3.0 (102-126-400)

Procedure & checklist 102-126-400 describes a method for constructing SMRTbell libraries using SMRTbell prep kit 3.0 (SPK 3.0) and SMRTbell barcoded adapter plate 3.0 that are suitable for generating HiFi reads on the Sequel II and IIe Systems for AAV sequencing applications

Procedure & checklist contents

1. AAV DNA input requirements per SMRT Cell 8M.
2. Recommendations for AAV DNA extraction and QC.
3. Recommendations for thermal annealing of ssAAV vectors.
4. Enzymatic workflow steps for preparation of multiplexed AAV SMRTbell libraries using SMRTbell prep kit 3.0 (102-182-700) and SMRTbell barcoded adapter plate 3.0 (102-009-200).
5. Guidance for pooling barcoded AAV SMRTbell libraries for multiplexed sequencing on a single SMRT Cell 8M.



Preparing multiplexed AAV SMRTbell libraries using SMRTbell prep kit 3.0

Procedure & checklist

PacBio

Before you begin

This protocol describes how to prepare multiplexed AAV libraries for sequencing on the Sequel®, Sequel II and Sequel IIe Systems. Multiplexing achieved using SMRTbell Barcoded adapters.

Once barcoded, samples can be pooled and purified for sequencing on the Sequel, Sequel II and Sequel IIe Systems. A total of 1 µg of pooled AAV DNA is required for SMRTbell library preparation. The input requirements per individual AAV sample depends on the multiplex level and can range from 250 ng for a 4-plex, to 42 ng per sample for a 24-plex preparation.

Preparation of AAV DNA for SMRTbell library preparation

AAV DNA may have several structures or formats: single-stranded DNA (ss), double-stranded DNA with some non-complimentary regions (ds), or single-stranded self-complimentary (sc) DNA, with an inverted repeat in the middle. All structures generally have inverted terminal repeats (ITR) at the ends.

SMRTbell template formation requires double-stranded DNA with blunt ends onto which hairpin SMRTbell adapters can be ligated, following A-tailing. Creating appropriate ds input DNA requires extra steps if the sample contains a substantial ss region without ds blunt ends, as shown on the left in Figure 1, below.

If the sample contains self-complementary molecules that have only one site accessible for hairpin adapter ligation, as shown on the right in Figure 1, it is important to collect and analyze data appropriately. Non-complementary ds molecules also require additional analysis.

Figure 1: Structures of ss (single-stranded) and sc (self-complimentary) DNA molecules. The ssDNA structure at left requires a second strand (as shown in yellow) for SMRTbell formation.

PN 102-126-400 EA V1 APR2022

PacBio

PacBio Documentation (102-126-400)

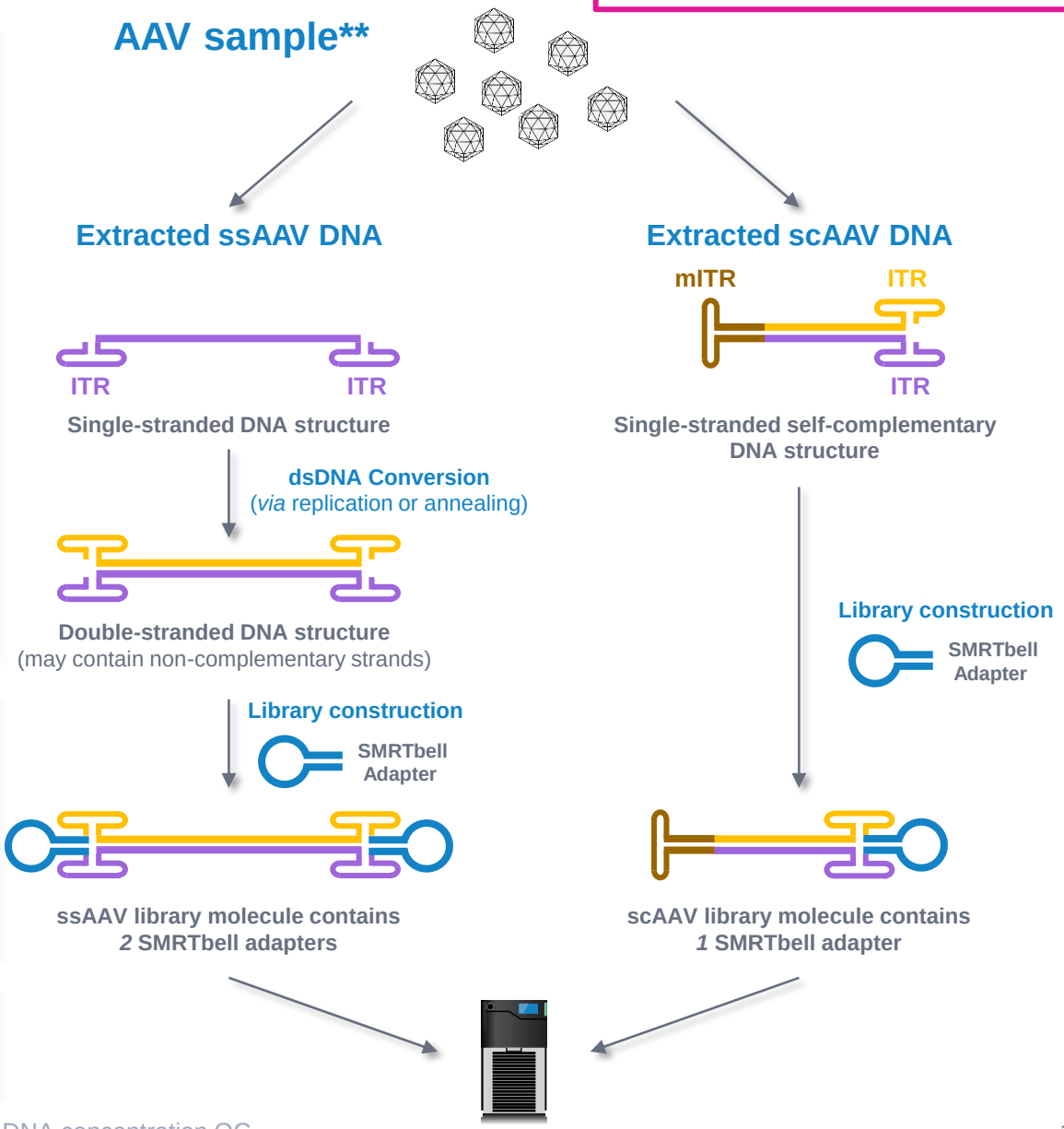
AAV SMRTbell construction workflow overview

**** Note:** Extracted AAV DNA samples typically contain a **mixture** of both ssAAV DNA and scAAV DNA

AAV workflow using SPK 3.0
(102-126-400)

Walk-away time	
10 min	DNAse I Treatment
45 min	AAV DNA extraction
	Safe Stop Point
75 min	Thermal annealing for ssAAV vectors
	Safe Stop Point — DNA QC (Sizing & Qubit)
20 min	Normalize DNA input & 1.3X SMRTbell bead cleanup
	Safe Stop Point
35 min	DNA repair & A-tailing
50 min	Adapter ligation & 1.3X SMRTbell bead cleanup
	Safe Stop Point
35 min	Nuclease treatment & 1.3X SMRTbell bead cleanup
	Safe Stop Point — DNA QC (Qubit)
20 min	Pool barcoded AAV samples & 1.3X SMRTbell bead concentration
	Safe Stop Point — DNA QC (Sizing & Qubit)
	Sequencing preparation & data collection

~5 hrs*



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

Workflow Step

- DNase I treatment
- AAV DNA extraction
- Thermal annealing for ssAAV vectors
- Normalize DNA input & cleanup
- Repair & A-tailing
- Adapter ligation & cleanup
- Nuclease treatment & cleanup
- Pool libraries & concentrate

DNase I treatment

DNase I treatment removes any non-encapsulated DNA from the vector

- We recommended using DNase I from NEB (M0303S) to treat extracted AAV DNA samples
- The capsid shell is resistant to DNase I treatment, therefore, DNase I will not degrade the encapsulated DNA.
- Add 20 units of DNase I to 7 x 10^{11} vector genomes (in 50 µL containing 1.5 x 10^{13} genome copies/mL)

Reagent	Volume (µL)
AAV vector	50
Nuclease-free Water	120
Reaction Buffer	20
DNase I (RNase-free)	10
Total Volume	**200**

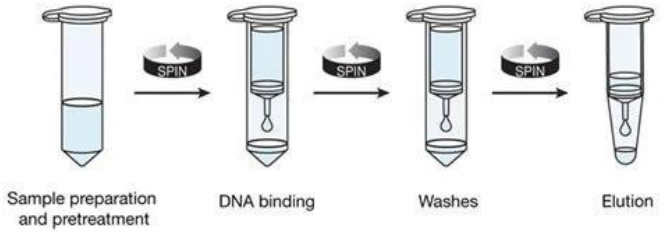
Workflow Step

- DNase I treatment
- AAV DNA extraction
- Thermal annealing for ssAAV vectors
- Normalize DNA input & cleanup
- Repair & A-tailing
- Adapter ligation & cleanup
- Nuclease treatment & cleanup
- Pool libraries & concentrate

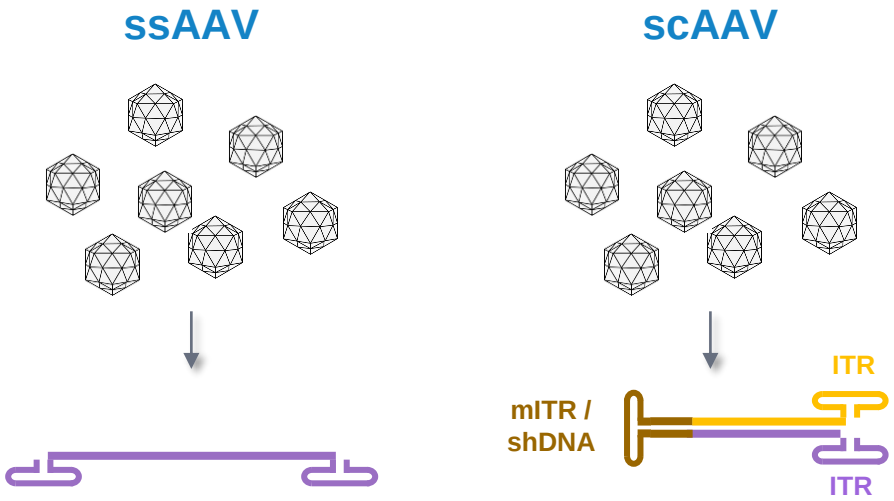
rAAV DNA extraction

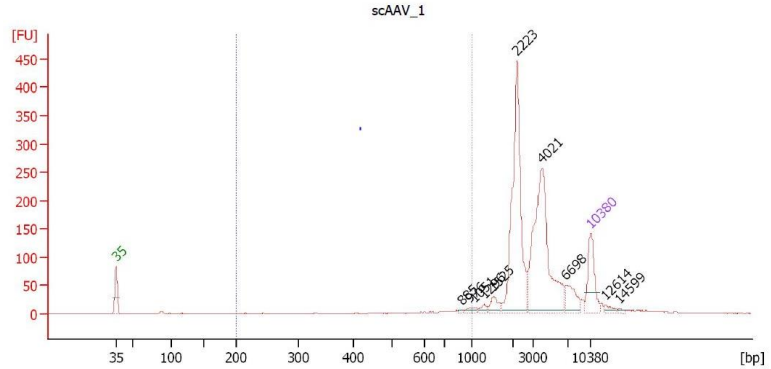
Extract DNA from DNase I treated sc/ss AAV vector by using the PureLink Viral RNA/DNA Mini Kit following the manufacturer's instructions

- PureLink Viral RNA/DNA Mini kit from Invitrogen (12280050) is recommended for AAV DNA extraction



- PureLink kit allows efficient lysis of viral particles at elevated temperatures and selective binding of viral nucleic acids to the silica matrix under highly denaturing condition (~1.5 hours)
- PureLink kit removes fragments <200 bp
  - If retention of DNA fragments <200 bp is desired, use an alternative extraction method such as phenol / chloroform / isoamyl alcohol (see Tran *et al*.)
- Perform DNA quantification QC on the extracted AAV DNA using a Qubit DNA assay kit and DNA sizing QC using a Bioanalyzer tool
- **Note:** Most extracted AAV DNA preparations typically contain a mixture of ssAAV and scAAV DNA





Example Bioanalyzer sizing QC analysis for an extracted scAAV DNA sample. The major peak at ~2 kb corresponds to the expected size range (1.8 kb) of this scAAV DNA sample.

Tran, N.T. et al. (2020) AAV-genome population sequencing of vectors packaging CRISPR components reveals design-influenced heterogeneity. Mol Ther Methods Clin Dev. 18:639–651.

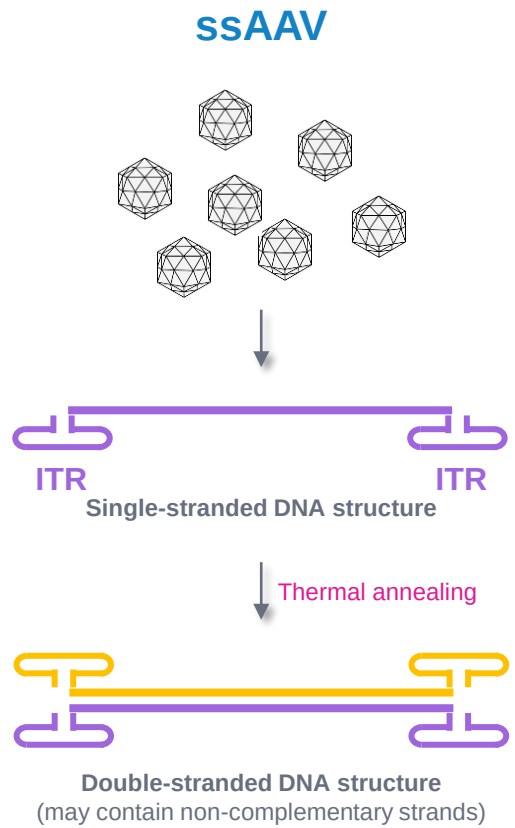
Workflow Step

- DNase I treatment
- AAV DNA extraction
- **Thermal annealing for ssAAV vectors**
- Normalize DNA input & cleanup
- Repair & A-tailing
- Adapter ligation & cleanup
- Nuclease treatment & cleanup
- Pool libraries & concentrate

Thermal annealing for ssAAV vectors

Perform thermal annealing to anneal + / – strands in ssAAV DNA to convert ssAAV vectors into a double-stranded DNA form suitable for SMRTbell library construction

- Thermal annealing step is only required for ssAAV vectors
  - Skip this step if working with scAAV vectors.
- To perform thermal annealing:
  - Add 1 vol. of ssAAV DNA sample to 9 vol. of 1X annealing buffer stock = 25 mM NaCl, 10 mM Tris-HCl [pH 8.5], 0.5 mM EDTA [pH 8]
  - Heat the diluted ssAAV DNA sample (prepared above) at 95°C for 5 min and then cool to 25°C (1 min for every −1°C) on a thermocycler
- Perform DNA quantification QC on the thermally annealed ssAAV DNA using a Qubit DNA assay kit and DNA sizing QC using a Bioanalyzer tool



if your sample contains substantial single-stranded DNA regions without double-stranded blunt ends as shown above for ssAAV, perform a thermal annealing step to create dsDNA suitable for SMRTbell library construction.

Tran, N.T. et al. (2020) AAV-genome population sequencing of vectors packaging CRISPR components reveals design-influenced heterogeneity. Mol Ther Methods Clin Dev. 18:639–651.

Workflow Step

- DNase I treatment
- AAV DNA extraction
- Thermal annealing for ssAAV vectors
- **Normalize DNA input & cleanup**
- Repair & A-tailing
- Adapter ligation & cleanup
- Nuclease treatment & cleanup
- Pool libraries & concentrate

Normalize DNA input & cleanup

Normalize the amount of AAV DNA for each sample to go into SMRTbell library construction to generate more even sequencing coverage across multiplexed samples

- Use a combined total of ≥1 μg of AAV DNA per SMRT Cell 8M
  - This is the combined total mass of samples that will be multiplexed
- Per sample input amounts should be 1 μg / number of samples
  - Use an equal mass of input DNA per AAV sample for multiplexing
  - Recommended per sample input AAV DNA amounts for different multiplex levels are shown in the table below

Multiplex level	DNA Input
4-plex	250 ng per sample
8-plex	125 ng per sample
12-plex	83 ng per sample
16-plex	63 ng per sample
24-plex	42 ng per sample

- Perform clean-up using 1.3X SMRTbell cleanup beads to concentrate the sample in an elution volume of 47 µL

PacBio

Workflow Step

- DNase I treatment
- AAV DNA extraction
- Thermal annealing for ssAAV vectors
- Normalize DNA input & cleanup
- **Repair & A-tailing**
- Adapter ligation & cleanup
- Nuclease treatment & cleanup
- Pool libraries & concentrate

Repair & A-tailing

Perform DNA damage repair, end repair and A-tailing reactions in a single tube

- For multiplexed preps:
  - Prepare a reagent master mix by adding the required components in the order and volume listed in the table below in a new tube
  - Adjust the buffer and enzyme component volumes for the total number of samples being prepared, plus 10% overage
- For non-multiplexed (individual) preps:
  - Add components directly to the sample from the previous step at the specified volumes and skip the preparation of a reagent master mix

✔	Tube	Reagent	Volume
	Previous	Contents from previous step	46 µL
	Purple	Repair buffer	8 µL
	Blue	End repair mix	4 µL
	Green	DNA repair mix	2 µL
		Total volume	**60 µL**

SMRTbell Prep Kit 3.0
(102-182-700)

- Incubate at 37ºC for 30 min followed by 65ºC for 5 min

PacBio

Workflow Step

- DNase I treatment
- AAV DNA extraction
- Thermal annealing for ssAAV vectors
- Normalize DNA input & cleanup
- Repair & A-tailing
- **Adapter ligation & cleanup**
- Nuclease treatment & cleanup
- Pool libraries & concentrate

Adapter ligation & cleanup

Ligate either barcoded or non-barcoded SMRTbell adapters to repaired and A-tailed AAV dsDNA to create SMRTbell templates

- For multiplexed preps:
  - Use barcoded adapters included with SMRTbell barcoded adapter plate 3.0 (102-009-200)

SMRTbell Barcoded Adapter Plate 3.0 (102-009-200) contains 96 ready-to-use barcoded SMRTbell adapters

- For non-multiplexed (individual) preps:
  - Use the standard (non-barcoded) adapter included with SMRTbell prep kit 3.0 (102-182-700)

SMRTbell Prep Kit 3.0 (102-182-700) contains the standard (non-barcoded) SMRTbell adapter

- Add 4 µL of adapter per sample, followed by the components listed in the table at right
- Incubate at 20ºC for 30 min
- Perform clean-up using 1.3X SMRTbell cleanup beads

Reaction Mix 2 (RM2)

✔	Tube	Component	Volume per Sample
	Yellow	Ligation mix	30 µL
	Red	Ligation enhancer	1 µL
		Total volume	**31 µL**

PacBio

Workflow Step

- DNase I treatment
- AAV DNA extraction
- Thermal annealing for ssAAV vectors
- Normalize DNA input & cleanup
- Repair & A-tailing
- Adapter ligation & cleanup
- **Nuclease treatment & cleanup**
- Pool libraries & concentrate

Nuclease treatment & cleanup

Perform a nuclease treatment step to reduce levels of undesirable reaction products that are not completely intact, circular SMRTbell templates

- For multiplexed preps:
  - Prepare a reagent master mix by adding the required components in the order and volume listed in the table below in a new tube
  - Adjust the buffer and enzyme component volumes for the total number of samples being prepared, plus 10% overage
- For non-multiplexed (individual) preps:
  - Add components directly to the sample from the previous step at the specified volumes and skip the preparation of a reagent master mix

Reaction Mix 3 (RM3)		
Tube	Component	Volume per sample
Light purple	Nuclease Buffer	5 µL
Light green	Nuclease Mix	5 µL
	Total Volume	**10 µL**

SMRTbell Prep Kit 3.0
(102-182-700)

- Incubate at 37ºC for 15 min
- Perform clean-up using 1.3X SMRTbell cleanup beads
- OPTIONAL: Take 1 µL and measure DNA concentration with the Qubit dsDNA HS assay kit to check for variable sample loss prior to pooling barcoded samples

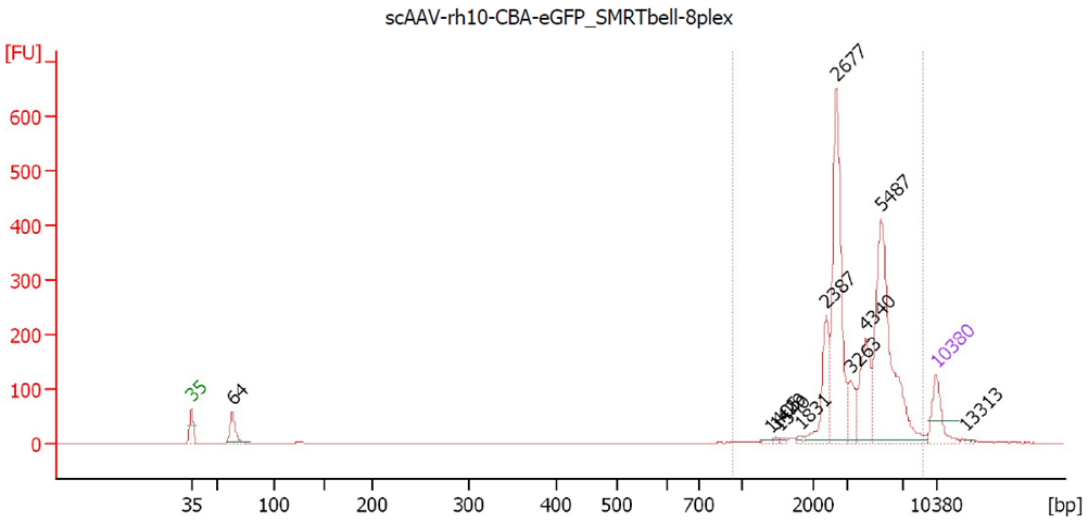
Workflow Step

- DNase I treatment
- AAV DNA extraction
- Thermal annealing for ssAAV vectors
- Normalize DNA input & cleanup
- Repair & A-tailing
- Adapter ligation & cleanup
- Nuclease treatment & cleanup
- **Pool libraries & concentrate**

Pool barcoded AAV samples & concentrate

Pool barcoded AAV SMRTbell library samples for sequencing on a single SMRT Cell

- Pooling libraries that have similar average insert sizes will provide more even sequence coverage between samples
  - Increasing differences in average molecular weights will lead to increasing differences in coverage across libraries because of differences in molarity for SMRT Cell loading
  - For sensitive experiments where balanced coverage is critical, consider pooling libraries in equal molar amounts
- 1.3X cleanup using SMRTbell cleanup beads to concentrate
- Perform DNA quantification QC on the final pooled AAV library using a Qubit DNA assay kit and DNA sizing QC using a Bioanalyzer tool

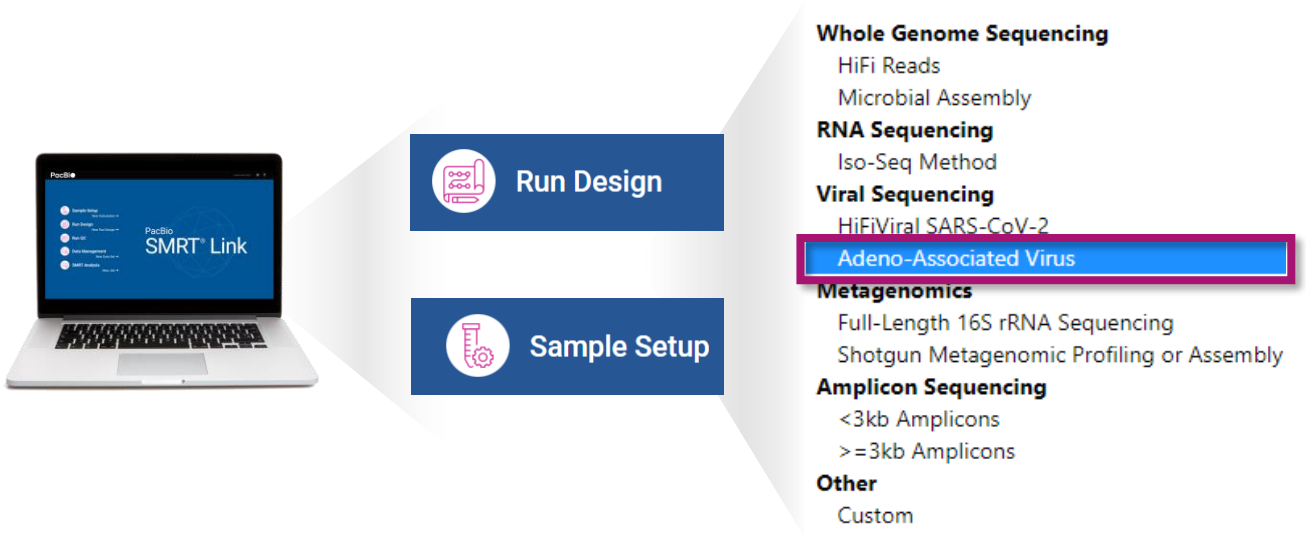


Example Bioanalyzer sizing QC analysis for a pooled 8-plex scAAV library sample. The major peak at ~2.6 kb corresponds to the expected size range (2.1 kb) of this scAAV DNA sample.

AAV library sequencing preparation workflow overview

Sample Setup & Run Design recommendations for AAV libraries

In SMRT Link Sample Setup & Run Design, select ‘Viral Sequencing’ / ‘Adeno-Associated Virus’ for application type



Binding kit 3.1 & cleanup beads (102-333-400) is recommended for preparing AAV samples for sequencing.

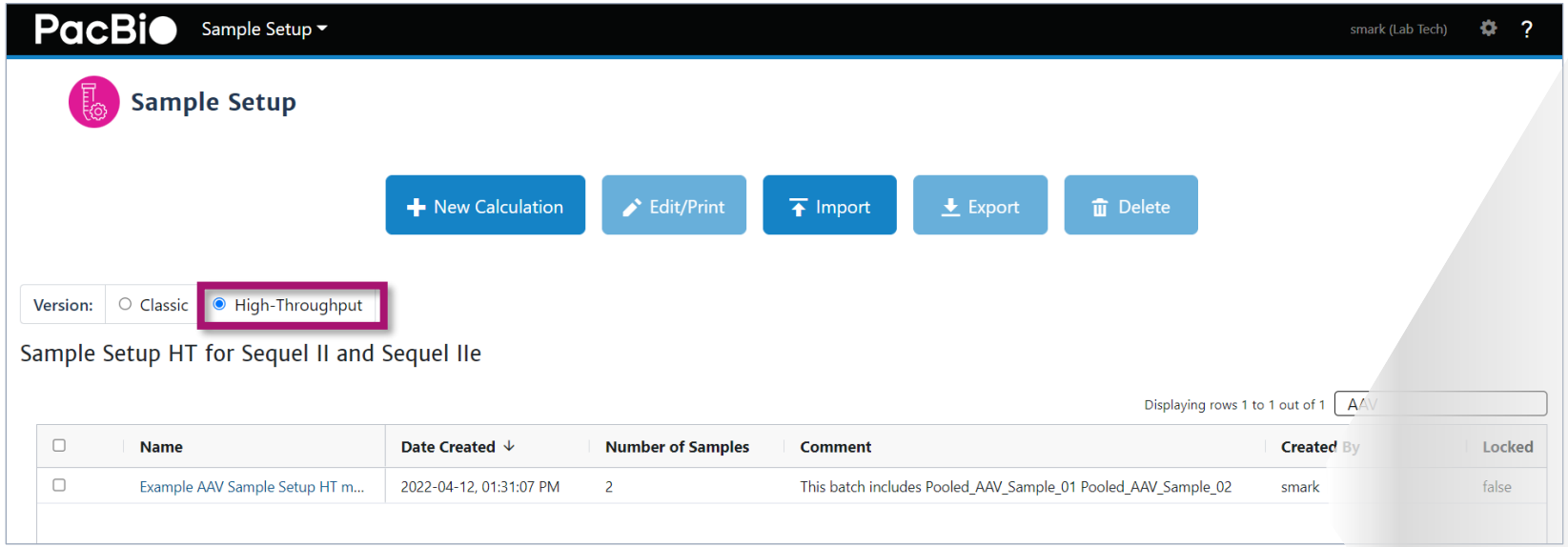
- We recommend using Sequel II binding kit 3.1 & cleanup beads (102-333-400) to perform ABC (anneal primer / bind polymerase / clean up complex) with AAV samples
- Refer to Quick reference card – Loading and pre-extension time recommendations for the Sequel II and IIe systems (101-769-100) for updates to ABC workflow for specific applications

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

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

AAV 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 AAV samples



- 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

Note: We recommend using an on-plate loading concentration (OPLC) of **70 – 150 pM** for AAV samples

	Sample Group
Actions	Copy / Remove / Lock / Automate
Name	Example AAV Sample Setup
Comment	This batch includes Pooled_AAV_Sample_01 Pooled_AAV_Sample_02
Application	Adeno-Associated Virus
Binding Kit	Sequel® II Binding Kit 3.1
Number of Samples	2 samples
SMRT Cells per Sample	1 cells
Available Volume per Sample	10 uL
Insert Size	3500 bp
Sample Concentration	12 ng/uL
Cleanup Anticipated Yield	60 %
Recommended Concentration on Plate	40-150 pM
Specify Concentration on Plate	90 pM
...imum Pipetting Volume	1 uL
Warnings	

Example Sample Setup HT mode worksheet for a batch comprised of two AAV samples.

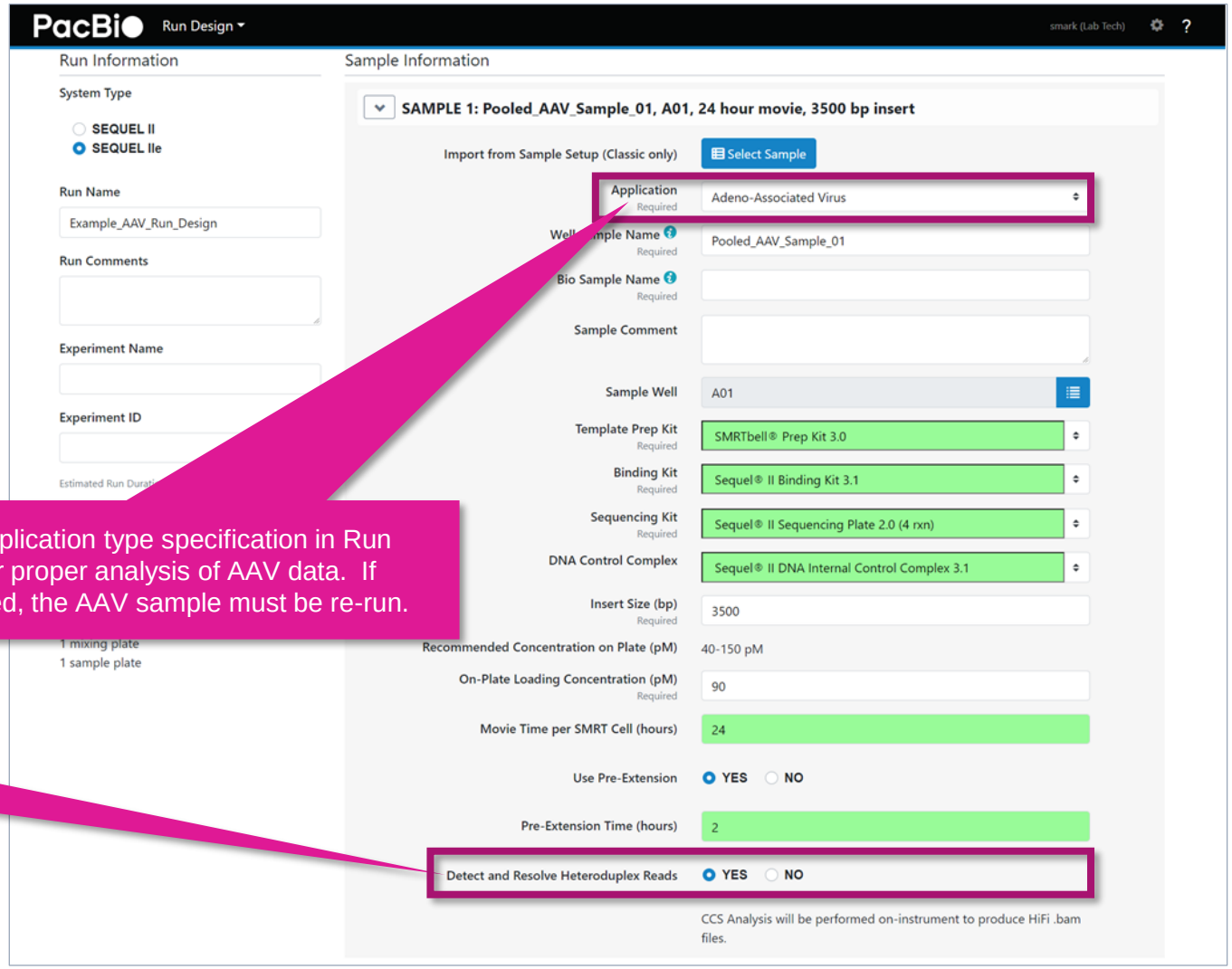
AAV Run Design guidance

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

- **IMPORTANT:** Select Adeno-Associated Virus from the Application field drop-down menu in SMRT Link Run Design
- Modified adapter calling + heteroduplex detection (HD) are automatically enabled by default when AAV application type is selected
- Specify YES for 'Detect and Resolve Heteroduplex Reads' regardless of whether you sample contains ssAAV, scAAV, or a mixture of ssAAV and scAAV DNA

IMPORTANT: AAV application type specification in Run Design is **required** for proper analysis of AAV data. If inadvertently not specified, the AAV sample must be re-run.

IMPORTANT: Heteroduplex detection is **required** for proper analysis of AAV data. If inadvertently not specified in Run Design, the AAV sample must be **re-run**.

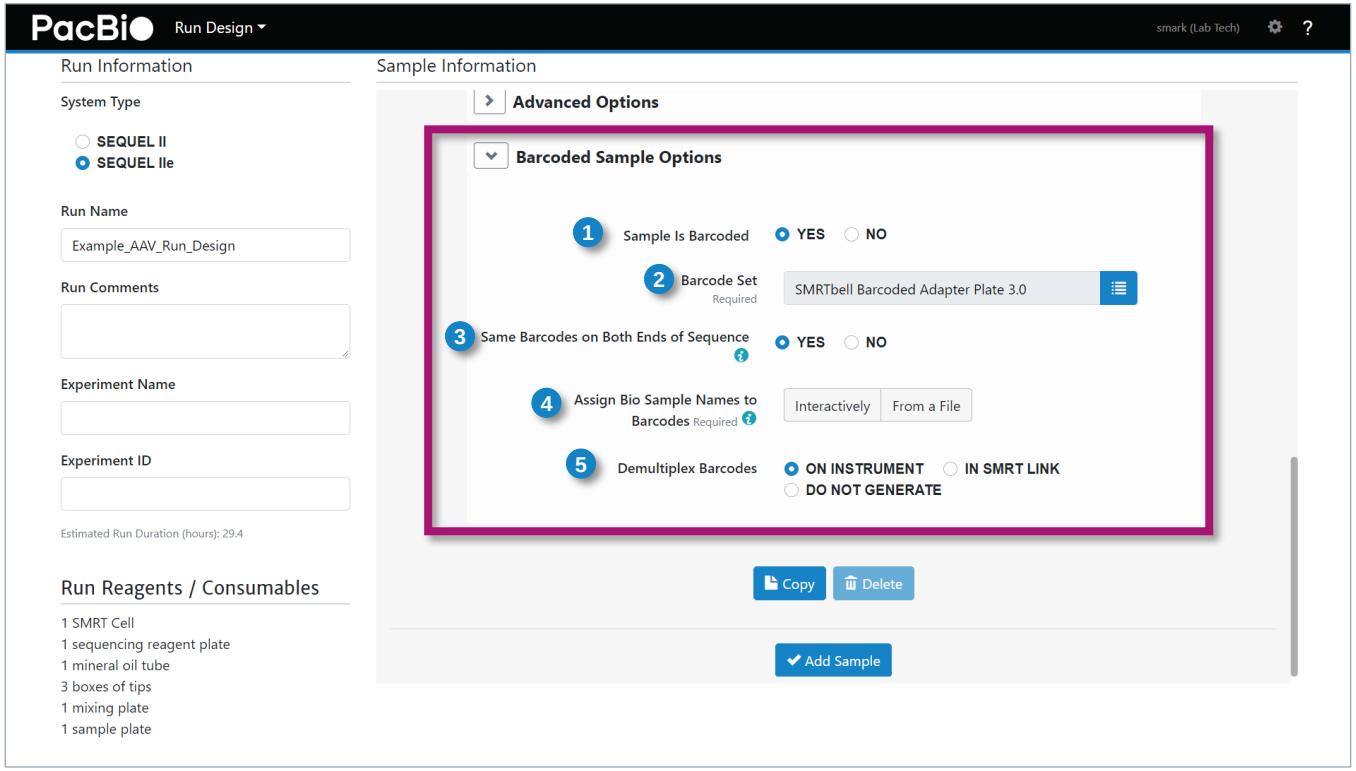


Example sample information entered into Run Design for sequencing an AAV sample.

AAV Run Design guidance (cont.)

OPTIONAL: Run Design setup procedure for automated barcode demultiplexing of pooled AAV library samples

1. Sample is Barcoded: YES
2. Barcode Set: Select 'SMRTbell barcoded adapter plate 3.0 (bc2001-bc2096)'
3. Same Barcodes on Both Ends of Sequence: YES
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.)

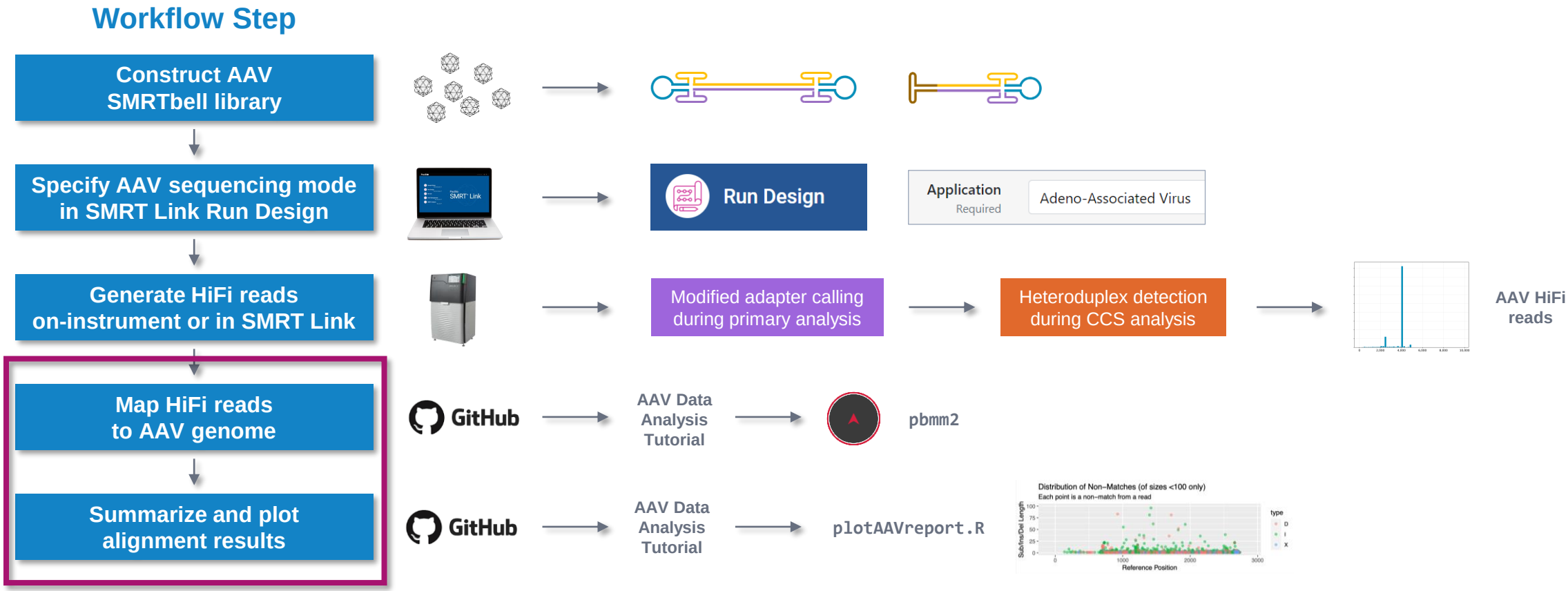


Example barcoding information entered into Run Design for sequencing a pooled AAV sample.

AAV data analysis workflow overview

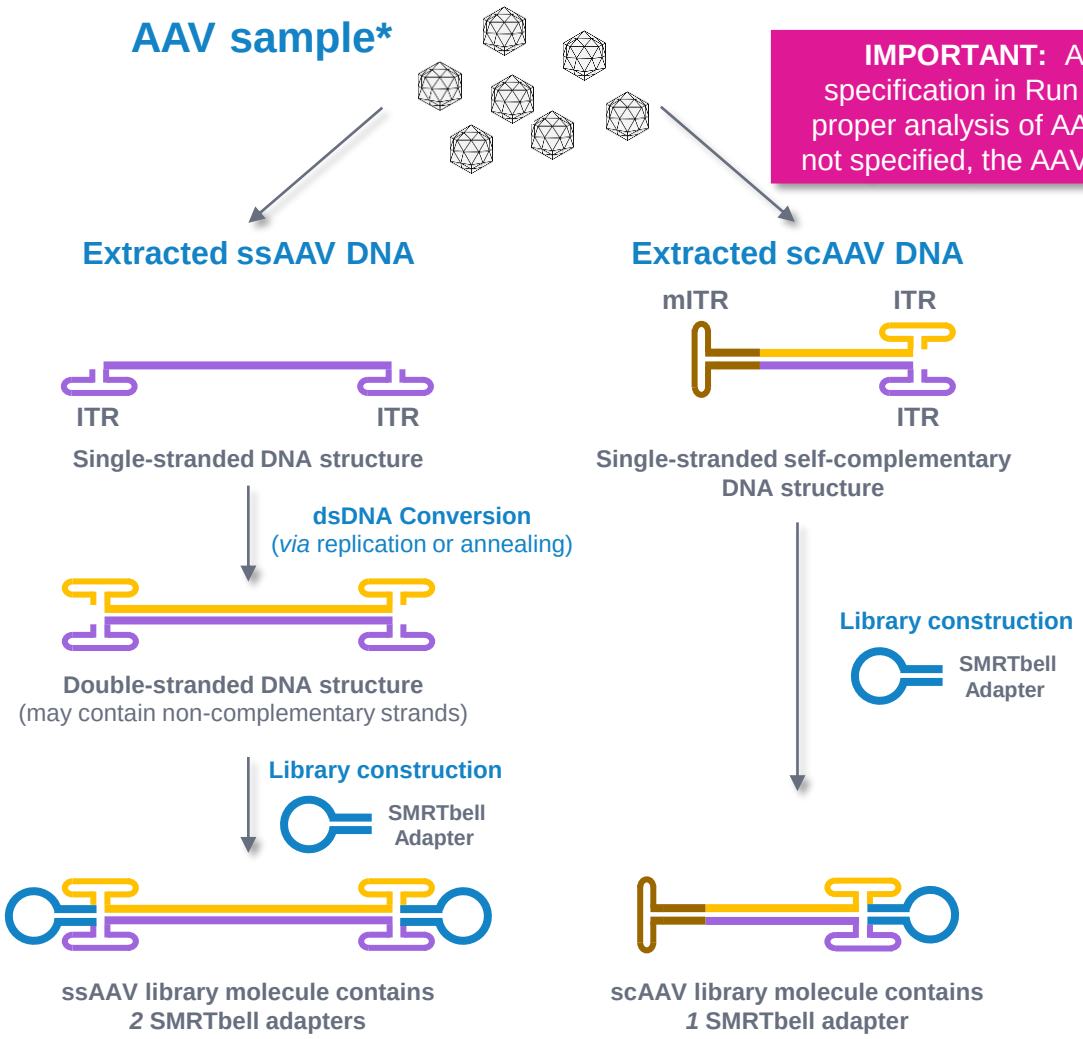
AAV data analysis workflow overview

After sequencing AAV samples using the AAV sequencing mode in Run Design, perform AAV data analysis QC using command line tools available in GitHub

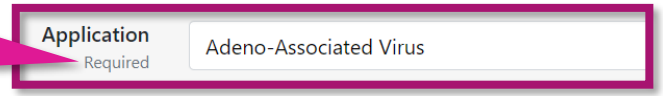


AAV application type specification in Run Design is IMPORTANT!

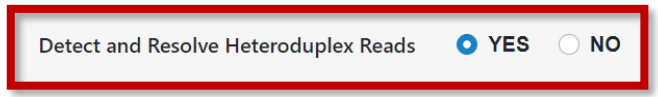
AAV sequencing mode in Run Design applies a modified adapter calling algorithm during post-primary analysis and enables automated heteroduplex (HD) read detection and resolution during CCS analysis



IMPORTANT: AAV application type specification in Run Design is **required** for proper analysis of AAV data. If inadvertently not specified, the AAV sample must be re-run.



Modified adapter calling + HD detection are **both** enabled by default when AAV application type is selected in Run Design

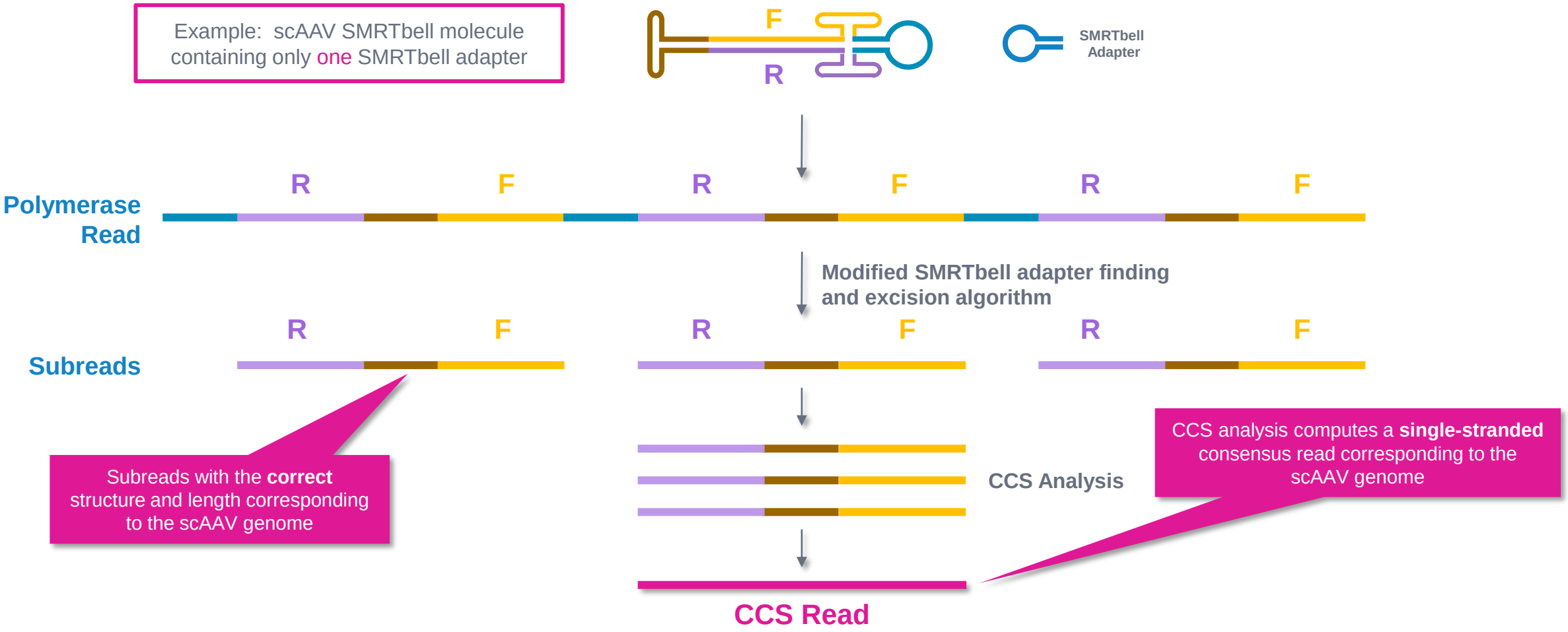


AAV data processing step	Purpose	Sequel IIe system workflow	Sequel II system workflow
Modified adapter calling	• Enables correct annotation of full-pass subreads within library molecules containing only 1 SMRTbell adapter • Required for analysis of scAAV libraries	Automated on-instrument analysis 	Automated on-instrument analysis 
Heteroduplex detection and splitting	• CCS analysis feature to identify a double-stranded sequence comprised of two non-complementary strands within the same library molecule • Separates the forward and reverse strands and generates a consensus read for each strand • Required for analysis of ssAAV libraries	Automated on-instrument analysis 	Automated in-SMRT Link analysis 

*** Note:** Extracted AAV DNA samples typically contain a mixture of ssAAV DNA and scAAV DNA

Modified adapter calling is automatically performed for AAV samples containing scAAV DNA (or a mixture of scAAV and ssAAV*)

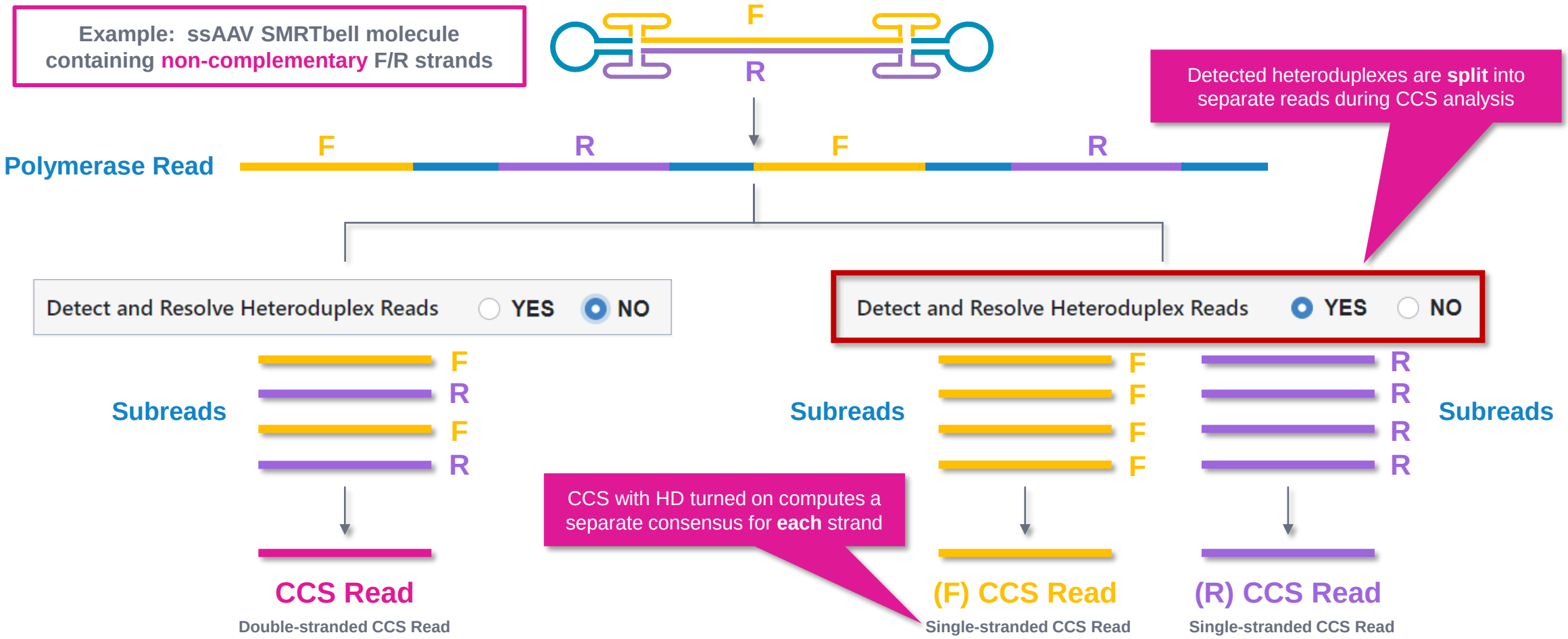
AAV sequencing mode in Run Design automatically applies a modified adapter calling algorithm to enable correct annotation of full-pass subreads for library molecules containing only *one* SMRTbell adapter



* If the AAV sample contains a mixture of scAAV and ssAAV, the ssAAV library molecules (containing *two* SMRTbell adapters) will undergo normal adapter calling during primary analysis.

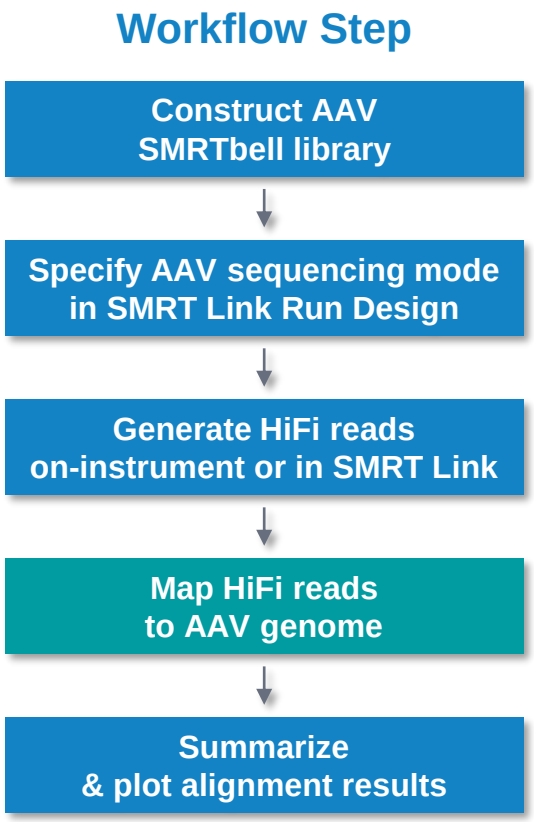
Heteroduplex detection and splitting is automatically performed for AAV samples containing ssAAV DNA (or a mixture of ssAAV and scAAV)

AAV sequencing mode in Run Design automatically applies a heteroduplex detection (HD) and splitting algorithm during CCS analysis of library molecules containing non-complementary regions



AAV data analysis QC recommendations

Map AAV HiFi reads to AAV reference genome using `pbmm2` command line tool available in GitHub



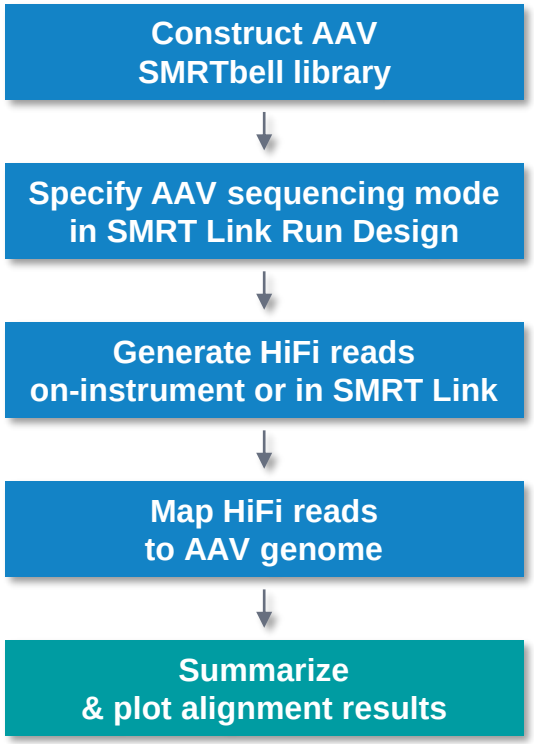
Map HiFi Reads to AAV Genome

- Example `pbmm2` command:
  - `pbmm2 –preset isoseq –sort –j <threads> <aav_genome.fa> <reads.bam> <mapped.sorted.bam>`
  - We use the `isoseq` preset to capture large deletion events (denoted by N in CIGAR)

AAV data analysis QC recommendations (cont.)

Summarize and visualize AAV alignment data using `summarize_AAV_alignment.py` and `plotAAVreport.R` command line tools available in GitHub

Workflow Step



Summarize & plot alignment results

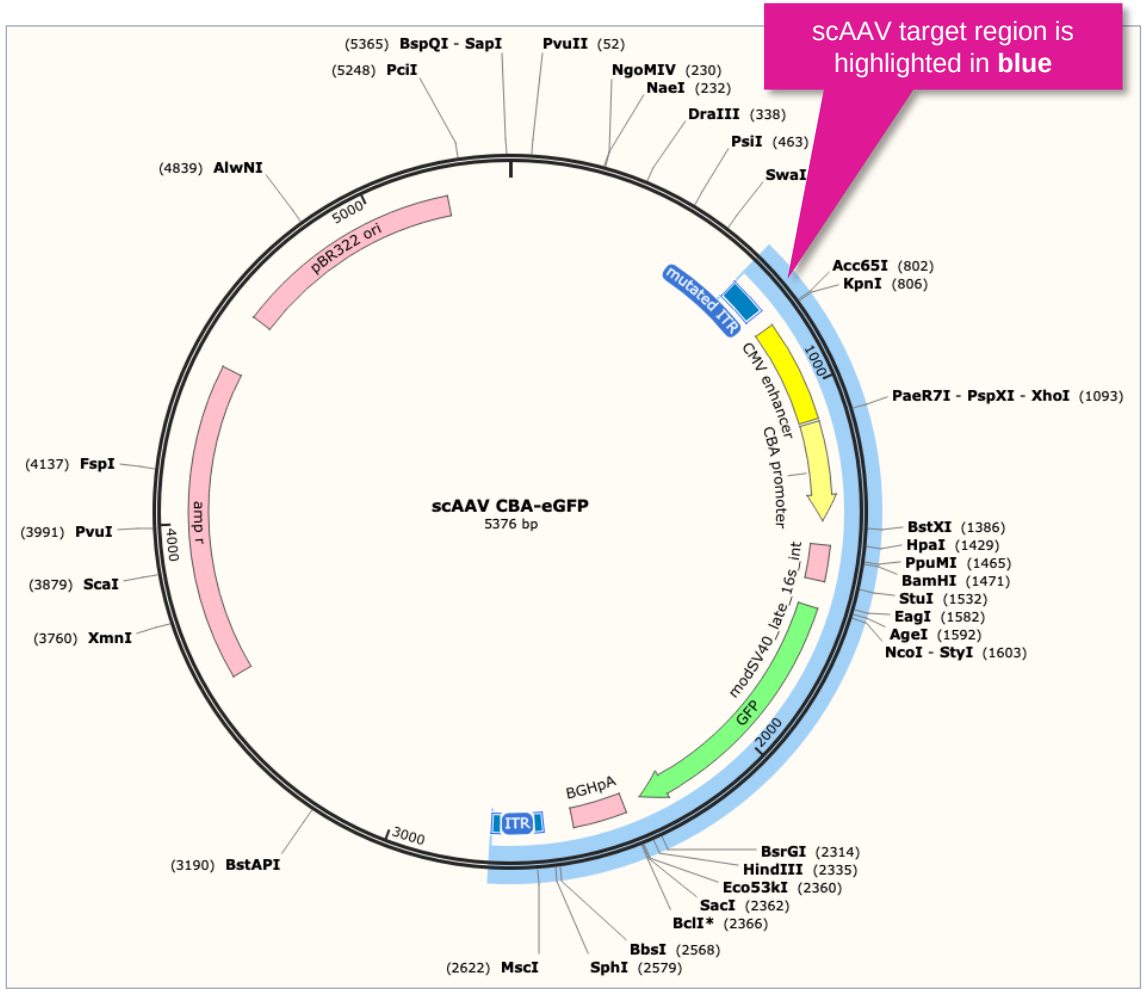
- After alignment, `summarize_AAV_alignment.py` is used to generate 3 csv files that give you information at varying levels of detail
  - `<prefix>.summary.csv`: mapping summary
    - Read name/length, if the read is mapped, if the alignment is supplementary, mapping coords, and mapping identity
  - `<prefix>.per_read.csv`: AAV type assignments
    - Read name/length, if that read has a primary/supp alignment, the assigned AAV type (sc, ss, unknown), and subtype (full length or partial)
  - `<prefix>.nonmatch_stat.csv`: breakdown of mistakes found in each read
    - Read name, error position, type of error (I, D, X), length of error
- Plot the alignment results using `plotAAVreport.R`
  - <prefix>. _AAV_report.pdf report
    - Distribution of mapped reference lengths, mapped identity to reference, fraction of reads that map to the reference start/end
    - Distribution of non−matches (deletions, insertions, substitutions) by reference position
    - Distribution of read lengths by assigned AAV type (ssAAV, scAAV, unknown)

AAV library example performance data

Example library QC and sequencing performance for AAV libraries

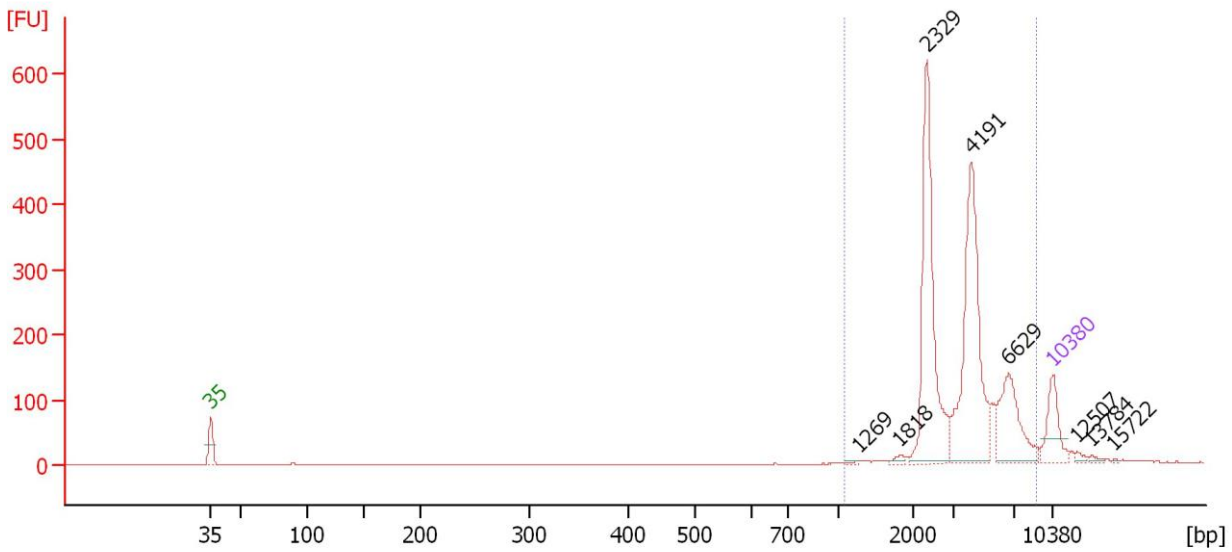
scAAV DNA extraction QC metrics

scAAV-CBA-eGFP Plasmid Map



scAAV target region: 662:2739 (2077 bp) (scAAV-CBA-eGFP, Vector Biolabs)

scAAV DNA Extraction QC



Example Bioanalyzer sizing QC analysis for an extracted scAAV DNA sample. The major peak at ~2.3 kb corresponds to the expected size range (~2.1kb) of this scAAV species (scAAV-CBA-eGFP, Vector Biolabs)

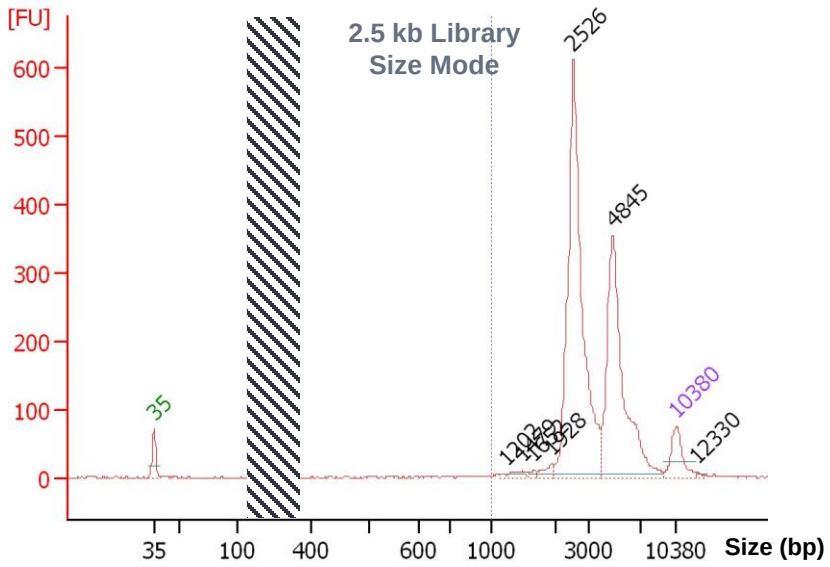
Example extracted DNA yield for an scAAV DNA sample extracted with the PureLink Viral RNA/DNA Mini kit from Invitrogen (12280050)

scAAV viral sample input	scAAV extracted DNA
1.5 x 10^{12} genome copies [1.5 x 10^{13} GC/mL (100µl)]	1000 ng

Example library QC and sequencing performance for AAV libraries (cont.)

SMRTbell library QC and primary sequencing metrics

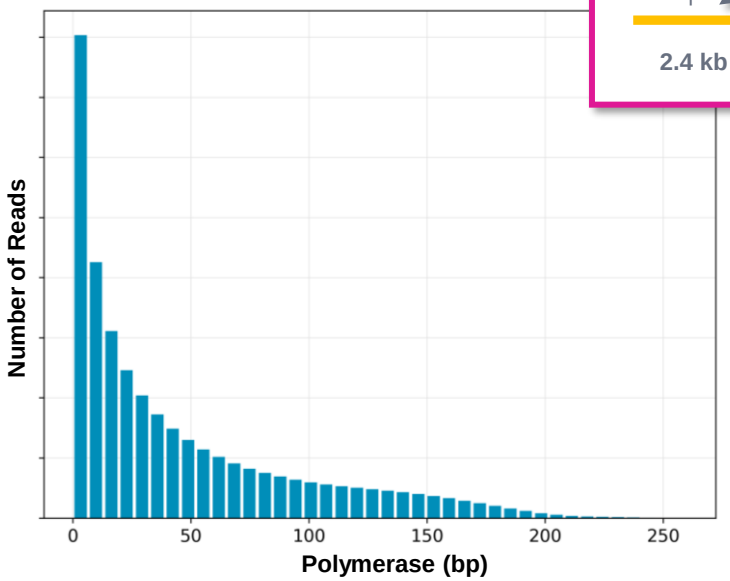
SMRTbell Library QC
(scAAV)



Extracted scAAV DNA input	1000 ng
Extracted scAAV DNA cleanup (%)	658 ng (65%)
Post-ligation & cleanup (%)	598 ng (59%)
Post-nuclease treatment & cleanup (%)	208 ng (20%)

Cleanup steps were performed using 1.3X SMRTbell cleanup beads.

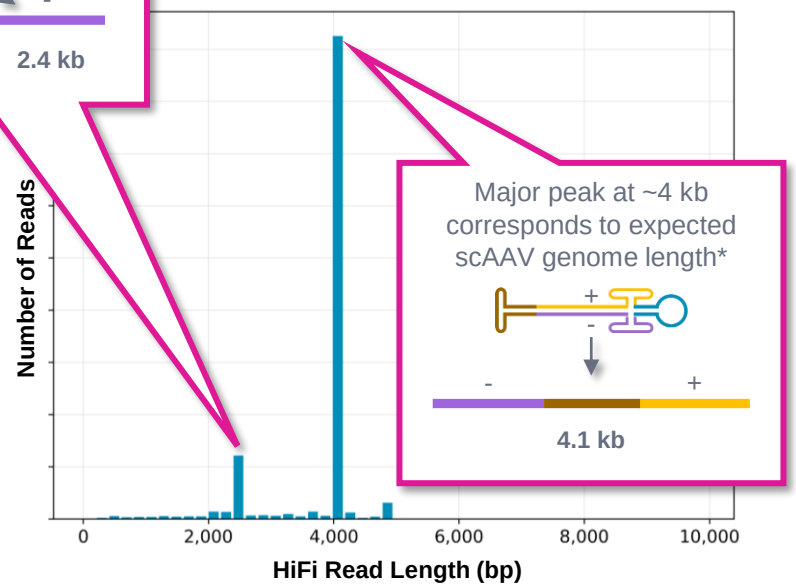
Raw Data Report
(scAAV)



Raw Base Yield	251.2 Gb
Mean Polymerase Read Length	46.0 kb
P0	30.6%
P1	68.1%
P2	1.3%

Example sequencing metrics for a scAAV sample run with Polymerase 2.1 / 90 pM on-plate concentration / 24-h movie time / 2-h Pre-Extension Time.

CCS Analysis Report
(1:10 ssAAV:scAAV Mix)



Mean HiFi Read Length	3,785 bp
Median HiFi Read Quality	Q32
HiFi Read Mean # of Passes	15

Total number of HiFi reads generated per SMRT Cell 8M is typically ≥2.5 Million for AAV samples achieve >50% *P1* loading.

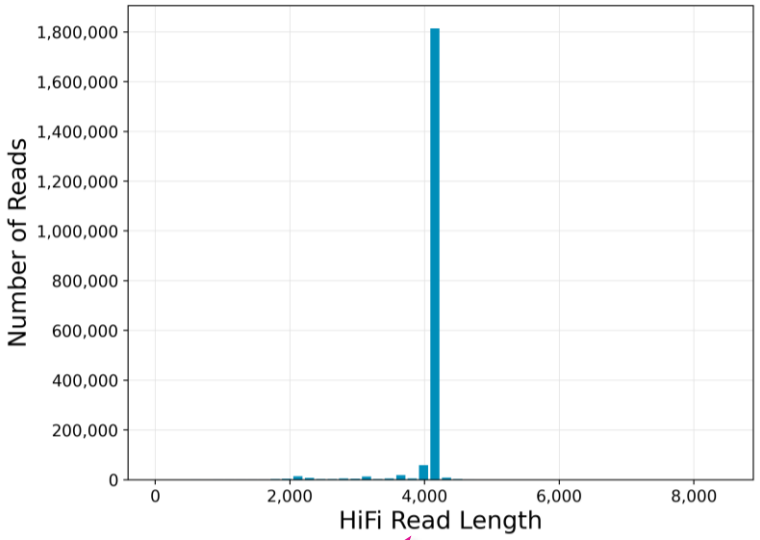
PacBio

* scAAV target length is 2077 bp (scAAV-CBA-eGFP, Vector Biolabs)

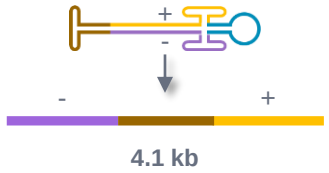
Example library QC and sequencing performance for AAV libraries (cont.)

Barcode demultiplexing metrics for a pooled (8-plex) scAAV sample

CCS Analysis Report



Major peak at ~4 kb corresponds to expected scAAV genome length*



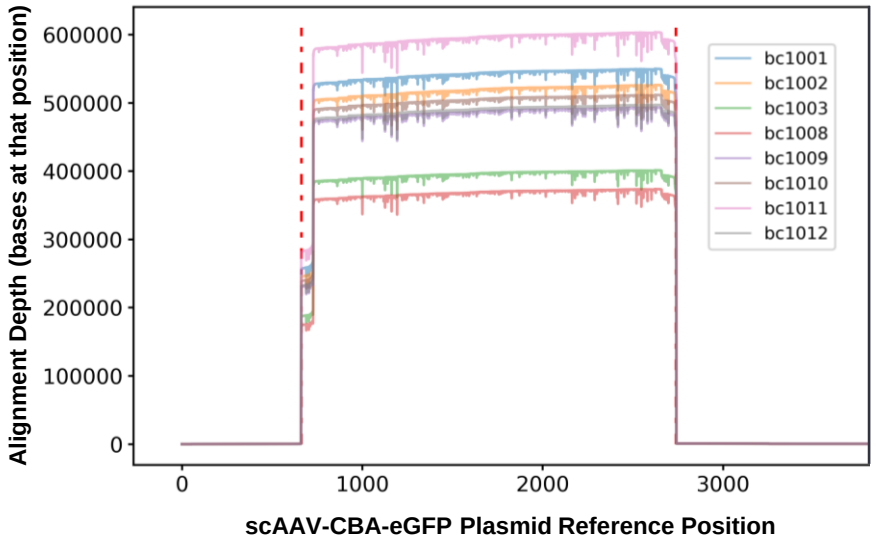
Demultiplexing Report

Barcode	Reads	Avg. Barcode Quality
bc1001	276,940	97.0
bc1002	264,996	97.0
bc1003	202,163	98.0
bc1008	188,205	97.0
bc1009	248,151	98.0
bc1010	257,343	98.0
bc1011	303,674	97.0
bc1012	250,317	98.0
None	854	0.0

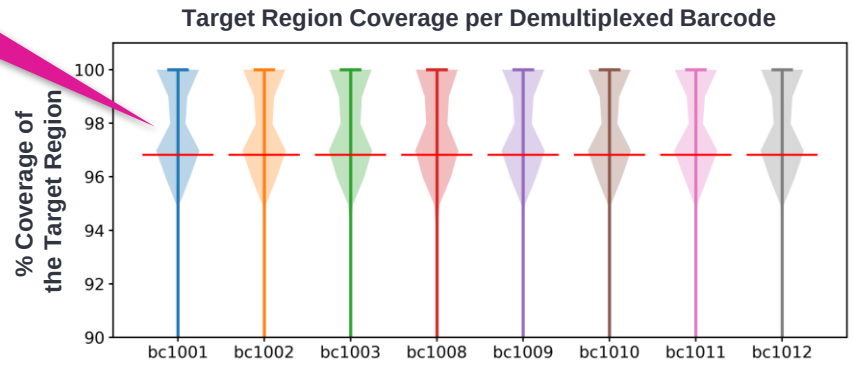
Report shows a similar yield of barcoded reads for each pooled AAV sample

97% median coverage was observed per barcoded AAV sample

Coverage Report for Demultiplexed scAAV Samples



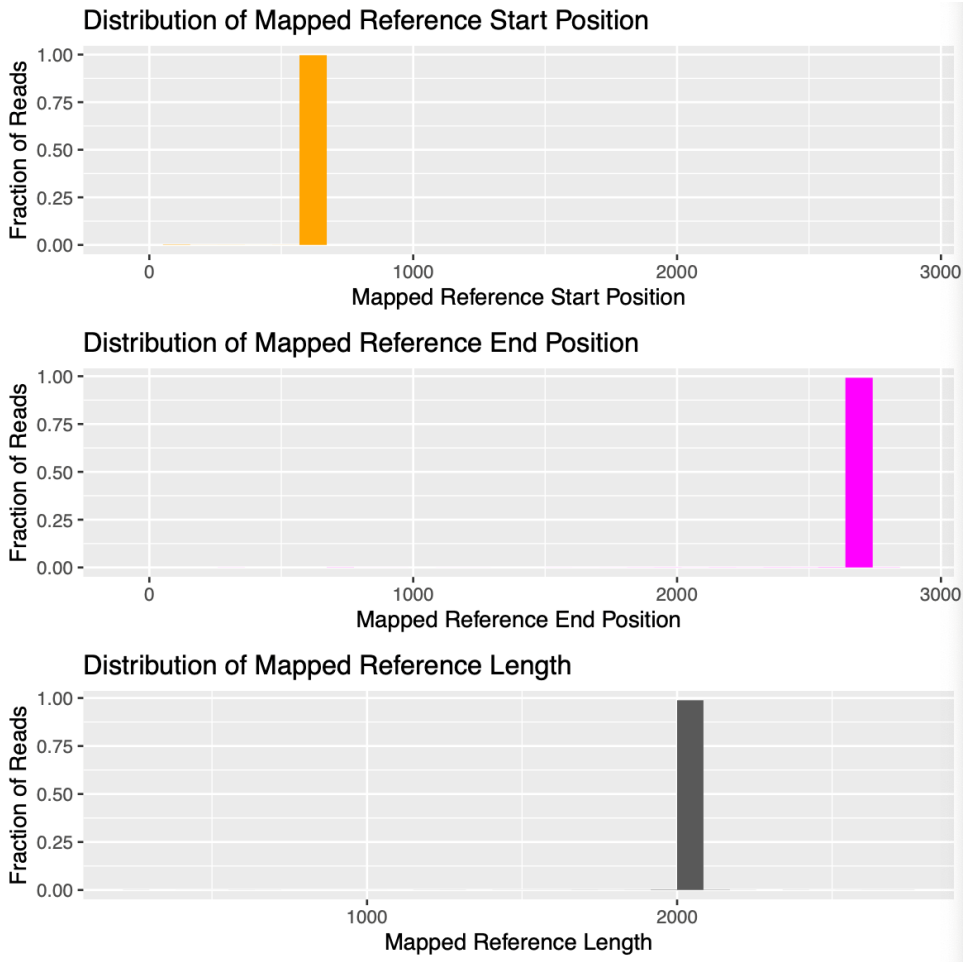
Plot of demultiplexed AAV HiFi reads aligned to the correct position of the reference transfer plasmid. scAAV target region: 662:2739 (2077 bp) (scAAV-CBA-eGFP, Vector Biolabs)



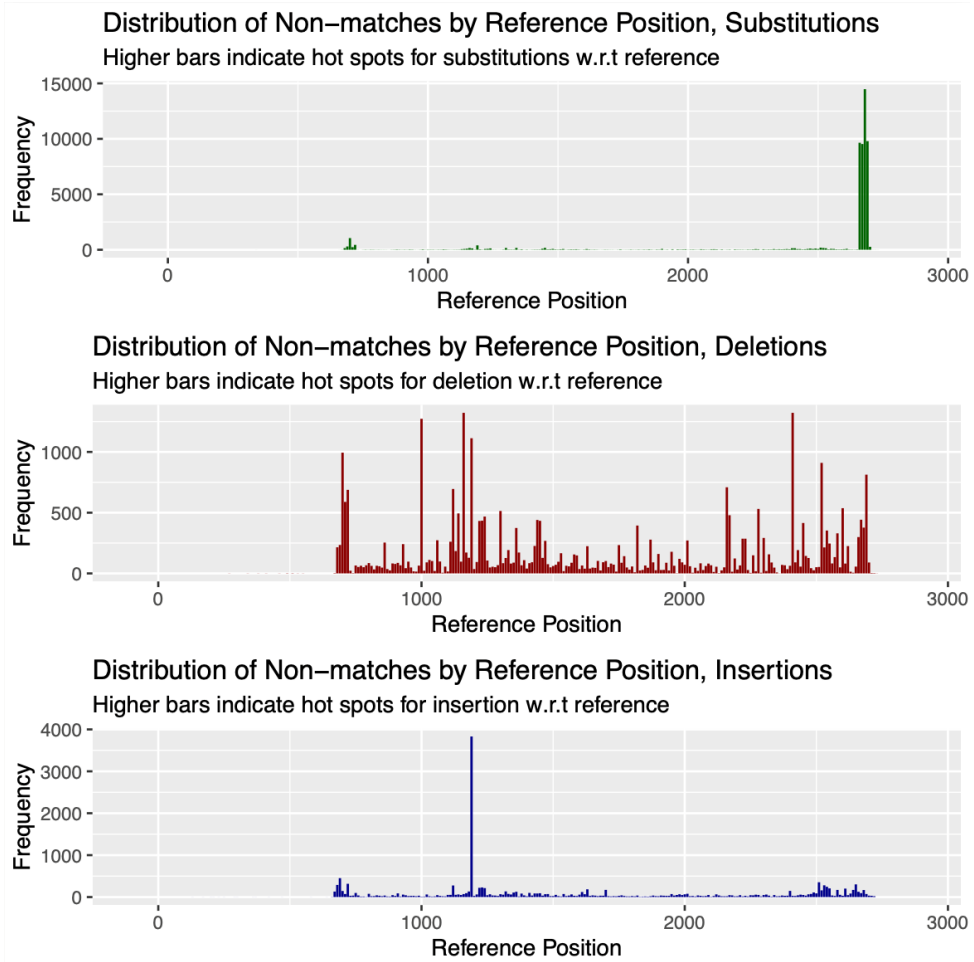
* scAAV target length is 2077 bp (scAAV-CBA-eGFP, Vector Biolabs)

Example plotAAVreport.R report output

Summarize and visualize AAV alignment data using `summarize_AAV_alignment.py` and `plotAAVreport.R` command line tools available in GitHub



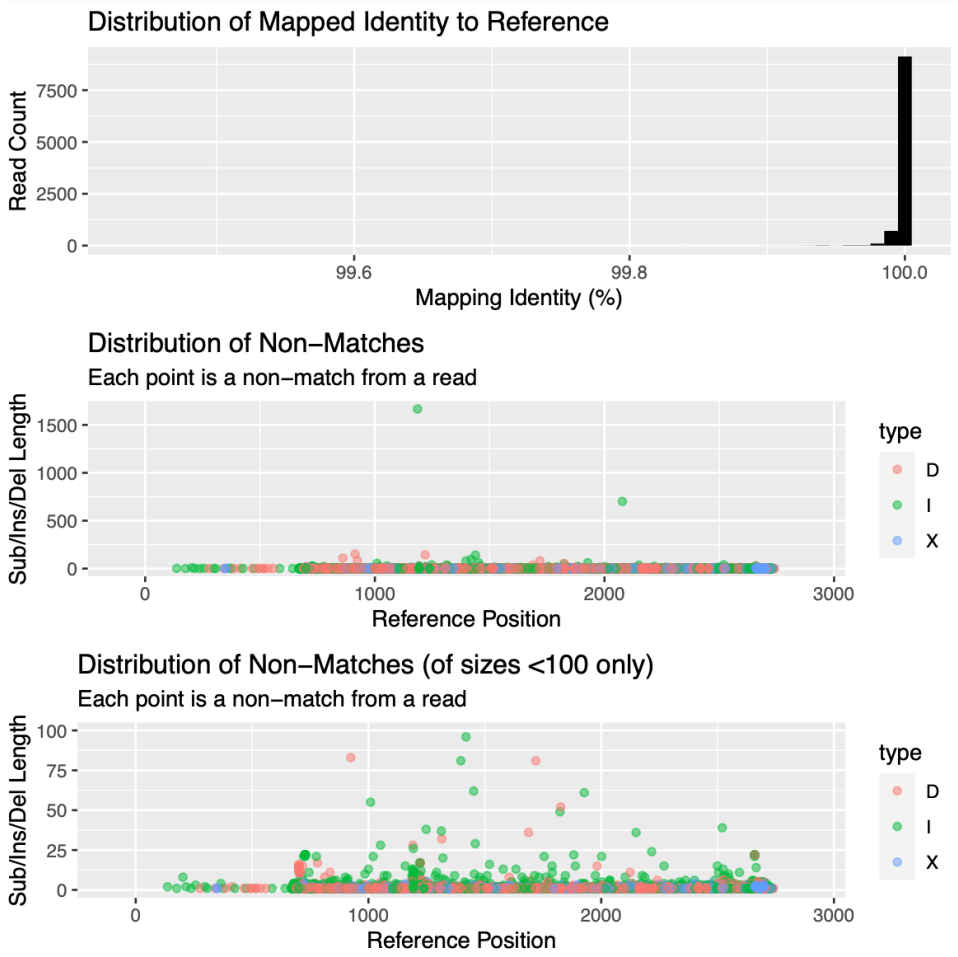
Distribution of mapped reference lengths and fraction of reads that map to the reference start/end for a scAAV control sample (scAAV-CBA-eGFP, Vector Biolabs).



Distribution of non−matches (deletions, insertions, substitutions) by reference position for a scAAV sample (scAAV-CBA-eGFP, Vector Biolabs).

Example plotAAVreport.R report output (cont.)

Summarize and visualize AAV alignment data using `summarize_AAV_alignment.py` and `plotAAVreport.R` command line tools available in GitHub



Distribution of mapped identities to reference and distribution of non−matches (deletions, insertions, substitutions) and their lengths by reference position for a scAAV sample (scAAV-CBA-eGFP, Vector Biolabs).

Err Type	Err Length	Count	Frequency (%)
D	1–10	33123	32.14
D	11–100	32	0.03
D	100–500	3	0.00
I	1–10	16532	16.04
I	11–100	66	0.06
I	100–500	1	0.00
I	>500	2	0.00
X	1–10	53289	51.71

Assigned Type	Assigned Subtype	Count	Frequency (%)
ssAAV	full	8	0.08
ssAAV	partial	8462	84.61
unknown	full	9	0.09
unknown	partial	1522	15.22

Frequency of non−matches (deletions, insertions, substitutions) by length [top] and assigned AAV type (ssAAV, scAAV or unknown) [bottom] for a ssAAV sample (pAV-CMV-GFP).

Technical documentation & applications support resources

Subtitle

Technical resources for AAV library preparation, sequencing & data analysis

AAV DNA extraction literature

- Tran, N.T. et al. (2020) AAV-genome population sequencing of vectors packaging CRISPR components reveals design-influenced heterogeneity. Molecular Therapy Methods & Clinical Development. 18:639 – 651. [Link]
- Guerin, K. et al. (2020) A novel next-generation sequencing and analysis platform to assess the identity of recombinant adeno-associated viral preparations from viral DNA extracts. Human Gene Therapy. 31:664 – 678 [Link]
- Lecompte et al. (2015) Advanced characterization of DNA molecules in rAAV vector preparations by single-stranded virus next-generation sequencing. Molecular Therapy Nucleic Acids. 4:E260. [Link]
- Gao, G. and Sena-Esteves, M. (2012). Introducing genes into mammalian cells: Viral vectors. Molecular cloning: A laboratory manual, Volume 2 New York: Cold Spring Harbor Laboratory Press. Pp. 1209 – 1313.

Sample preparation literature

- Procedure & checklist – Preparing multiplexed AAV SMRTbell libraries using SMRTbell prep kit 3.0 (102-126-400)
- Quick reference card – Loading and pre-extension recommendations for the Sequel II and IIe systems (101-769-100)
- Overview – Sequel systems application options and sequencing recommendations (101-851-300)
- Technical note: Preparing DNA for PacBio HiFi sequencing – Extraction and quality control (101-061920)
- Technical overview: AAV library preparation using SMRTbell prep kit 3.0 (102-390-400)

Technical resources for AAV library preparation, sequencing & data analysis (cont.)

Data analysis resources

- Tutorial: Analyzing AAV data (GitHub): https://github.com/Magdoll/AAV/wiki/Tutorial:-Analyzing-AAV-Data
- SMRT Link v11.0 user guide (102-278-200)
- SMRT Tools v11.0 reference guide (102-278-500)
- Sequel II and IIe systems: Data files (102-144-100)

Example PacBio data sets

AAV sequencing application	Dataset	Data type	PacBio system
scAAV (8-plex)	2021-scAAV/rh10-CBA-eGFP	HiFi Reads	Sequel II System
ssAAV (Single sample)	2021-pAV-CMV-GFP	HiFi Reads	Sequel II System

Posters

- Ranade, S. et al. (2018). High-throughput SMRT sequencing of clinically relevant targets. [Link]

Technical resources for AAV library preparation, sequencing & data analysis (cont.)

Publications

- Tran, N.T. et al. (2020) AAV-genome population sequencing of vectors packaging CRISPR components reveals design-influenced heterogeneity. Molecular Therapy Methods & Clinical Development. 18:639 – 651. [Link]
- Tai, P.W. et al. (2018) Adeno-associated virus genome population sequencing achieves full vector genome resolution and reveals human-vector chimeras. Molecular Therapy Methods & Clinical Development. 9:130–141. [Link]
- Paulk, N.K. et al. (2018) Bioengineered AAV Capsids with Combined High Human Liver Transduction In Vivo and Unique Humoral Seroreactivity. Molecular Therapy. 26:289 – 303. [Link]
- Xie, J. et al., 2017. Short DNA hairpins compromise recombinant adeno-associated virus genome homogeneity. Molecular Therapy. 25(6):1363–1374. [Link]
- Hüser, D. et al. (2014) Adeno-associated virus type 2 wild-type and vector-mediated genomic integration profiles of human diploid fibroblasts analyzed by third-generation PacBio DNA sequencing. Journal of virology, 88:11253–11263.[Link]

Webinars

- PacBio Journal Club (2021): AAV-genome population sequencing of vectors packaging CRISPR components reveals design-influenced heterogeneity [Link]
- Genewiz Cell and Gene Therapy Treatments Event (2021): Gene editing validation with single-molecule resolution using highly accurate HiFi reads [Link]
- PacBio Webinar (2019): Highly accurate SMRT sequencing for gene editing applications [Link]

www.pacb.com