

# Technical overview – Kinnex library preparation using Kinnex full-length RNA kit

Sequel II and IIe systems ICS v11.0 Revio system ICS v13.0 SMRT Link v13.0

PN 103-344-700 Rev 01 | January 2024

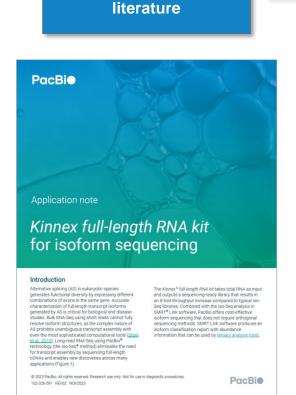
# Kinnex library preparation using Kinnex full-length RNA kit

#### **Technical Overview**

- 1. Kinnex full-length RNA method overview
- 2. Kinnex full-length RNA library preparation workflow details
- 3. Kinnex full-length RNA sequencing preparation workflow details
- 4. Kinnex full-length RNA example sequencing performance data
- 5. Kinnex full-length RNA data analysis workflow overview
- 6. Technical documentation & applications support resources



# Kinnex library preparation using Kinnex full-length RNA kit: Getting started



**Application-specific** 

#### Application note - Kinnex full-length RNA kit for isoform sequencing (102-326-591)

Summary overview of application-specific library preparation and data analysis workflow recommendations.

**Application-specific** protocol

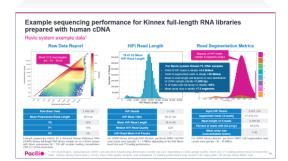
PacBi● Preparing Kinnex™ libraries using the Kinnex full-length RNA kit Before you begin This procedure describes the workflow for constructing Kinnex full-length RNA libraries from total RNA samples for Up to 4 SMRT Cells for Sequel II/IIe sy RIN (RNA integrity number) ≥7.0 300 ng per library (minimum concentration 43 ng/uL per li © 2023 PacBio, All rights reserved. Research use only. Not for use in diagnostic procedures.

#### **Procedure & checklist - Preparing Kinnex** libraries using the Kinnex full-length RNA kit (103-238-700)

Technical documentation containing applicationspecific library preparation protocol details.

**Application-specific** technical overview





#### **Technical Overview - Kinnex library** preparation using Kinnex full-length RNA kit (103-344-700)

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



#### cDNA synthesis & amplification

(Iso-Seg express 2.0 kit)

Library preparation,

sequencing & analysis

300 ng input total RNA per sample RIN (RNA integrity number) ≥7.0 Sample multiplexing options available



#### **Kinnex library preparation**

(Kinnex PCR 8-fold kit + Kinnex concatenation kit)

Use amplified cDNA to generate Kinnex library containing 8-segment array Sample multiplexing up to 48-plex supported



#### **SMRT** sequencing

(Seguel II/IIe & Revio systems)

Perform ABC\* and sequence Kinnex libraries on PacBio long-read systems



#### Data analysis (SMRT Link)

Use SMRT Link Read Segmentation data utility to split arrayed transcript HiFi reads

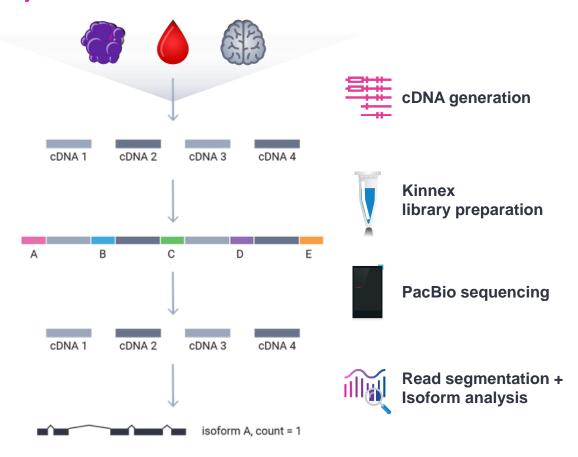
Use SMRT Link Iso-Seg analysis application to identify novel genes and isoforms with abundance information





# Kinnex full-length RNA method overview

Use Kinnex full-length RNA kit to perform high-accuracy, full-length isoform sequencing with PacBio long-read systems



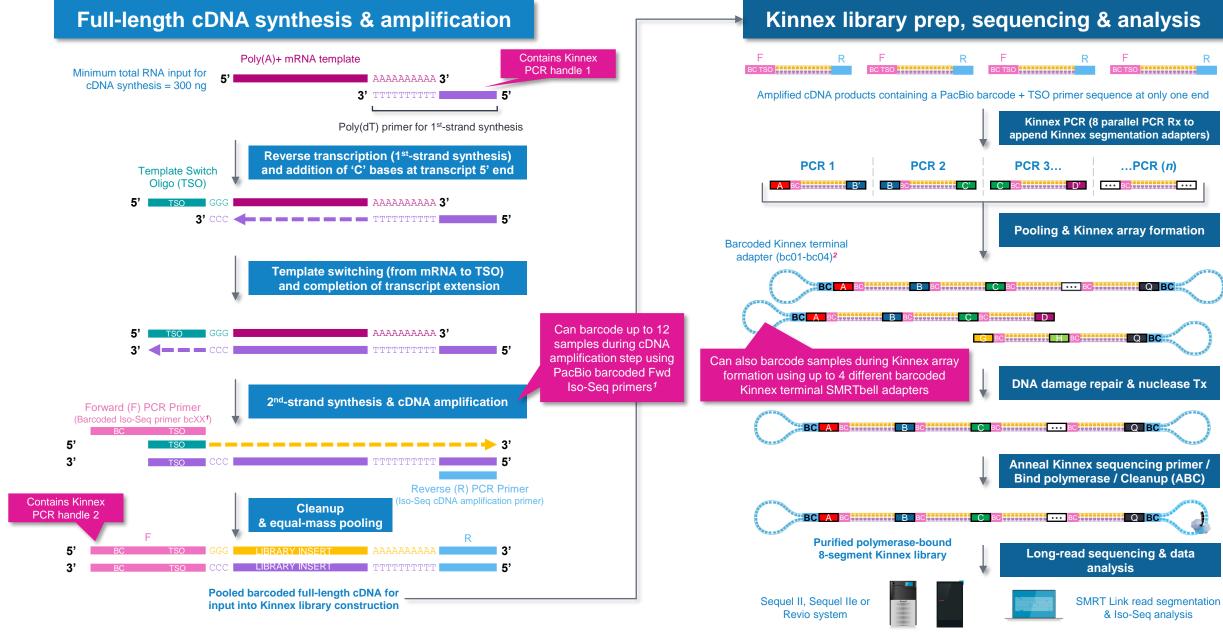
isoform B, count = 10

isoform C, count = 5

- Input 300 ng total RNA, RIN ≥7
- Generate up to 12-plex barcoded cDNA using Iso-Seq express 2.0 kit (103-071-500)
- 2-day Kinnex library preparation using Kinnex full-length RNA kit (103-072-000)
- SMRT Link Run Design support for 'Kinnex full-length RNA' application type with auto-analysis (read segmentation + isoform analysis)
- SMRT Link Iso-Seq isoform-classification software to identify novel genes and isoforms with abundance information



# Kinnex full-length RNA method overview (cont.)



<sup>1</sup> Twelve barcoded Iso-Seq primers (Iso-Seq primer bc01–12) are available for cDNA amplification step.

<sup>&</sup>lt;sup>2</sup> Kinnex adapter barcode sequences can be downloaded from <u>SMRT Link</u> Data Management module.

# Kinnex full-length RNA library preparation procedure description

Procedure & checklist – Preparing Kinnex libraries using the Kinnex full-length RNA kit (103-238-700) describes the workflow for constructing Kinnex libraries from total RNA samples using the **Iso-Seq** express 2.0 kit and Kinnex full-length RNA kit for sequencing on PacBio Sequel II/IIe & Revio systems

Overview	
Samples	1-24
Workflow time	8 hours (for up to 24 samples)
Number of SMRT® Cells per Kinnex library Prep	Up to 2 SMRT Cells for Revio system Up to 4 SMRT Cells for Sequel II/IIe systems
RNA input	
Quality/size distribution	RIN (RNA integrity number) ≥7.0
Quantity	300 ng per library (minimum concentration 43 ng/µL per library)



Kinnex full-length RNA kit 103-072-000 (12 rxn)



103-071-500 (24 rxn)



Kinnex full-length RNA library template (~12–16 kb)
Contains 8 concatenated full-length cDNA cDNA segments



1 acbio <u>bocumentation</u> (103-230-700)

- Kinnex full-length RNA library prep protocol uses Kinnex full-length RNA kit and Iso-Seq express 2.0 kit
  - Do not use SMRTbell prep kit 3.0 with this protocol



# Kinnex full-length RNA kit bundle and Iso-Seq express 2.0 kit components

Kinnex full-length RNA kit bundle and Iso-Seq express 2.0 kit provide full support for Kinnex library prep workflow

#### **Iso-Seq express 2.0 kit (103-071-500)**

Includes Iso-Seq Express template switching oligo, barcoded cDNA PCR Primers, and other reagents needed for performing 1<sup>st</sup>-strand cDNA synthesis and PCR amplification of cDNA products generated from input total RNA.

Iso-Seq express 2.0 kit components				
Compo	nent	Description		
1		<ul> <li>Iso-Seq RT buffer</li> <li>For 1<sup>st</sup>-strand cDNA synthesis</li> </ul>		
2	area.	<ul> <li>Iso-Seq RT primer mix</li> <li>For 1<sup>st</sup>-strand cDNA synthesis</li> </ul>		
3		<ul> <li>Iso-Seq RT enzyme mix</li> <li>For 1<sup>st</sup>-strand cDNA synthesis</li> </ul>		
4		<ul> <li>Iso-Seq template switch oligo</li> <li>For 1<sup>st</sup>-strand cDNA synthesis</li> </ul>		
5		<ul><li>Iso-Seq cDNA PCR mix</li><li>Enzyme nucleotide mix for cDNA amplification</li></ul>		
6		<ul><li>Iso-Seq cDNA amplification primer</li><li>Reverse primer for cDNA amplification</li></ul>		
7		<ul> <li>Iso-Seq primers (bc01 – bc12)</li> <li>Barcoded forward primers (bc01 – bc04) for cDNA amplification</li> </ul>		

#### Kinnex full-length RNA kit bundle (103-072-000)

Includes Kinnex PCR kit, Kinnex concatenation and ancillary DNA cleanup reagents needed for incorporation of Kinnex segmentation adapters and Kinnex array formation for generating Kinnex full-length RNA libraries.

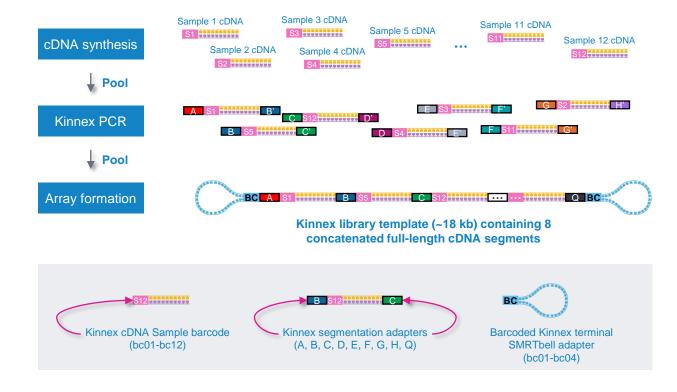
# Kinnex full-length RNA kit components Component **Description Kinnex PCR 8-fold kit (12 rxn)** Contains reagents for Kinnex PCR to incorporate segmentation adapters **Kinnex concatenation kit (12 rxn)** Contains reagents for Kinnex array formation and SMRTbell template construction Includes barcoded Kinnex adapter mixes (bc01 - bc04)**SMRTbell cleanup beads** For DNA cleanup **Elution buffer** For DNA cleanup

# Kinnex full-length RNA library barcoding options for sample multiplexing

Kinnex full-length RNA library preparation procedure supports up to 48-plex sample multiplexing

Kinnex full-length RNA library preparation procedure supports up to 48-plex sample multiplexing through combined use of:

- → 12 different barcoded cDNA amplification PCR primers (bc01 bc12)
- → 4 different barcoded Kinnex terminal SMRTbell adapters (bc01 bc04)





# Kinnex full-length RNA experimental design considerations

Kinnex full-length RNA application use case recommendations for PacBio systems

Example application	Human genetics disease studies	Biopharma for identifying highly expressed targets	Plant & animal whole genome annotation			
Experimental goal	Isoform discovery and quantification of moderate-to-rare transcripts	Isoform discovery of high expressed transcripts	Comprehensive transcript annotation in a species			
Example study design	Disease vs. normal tissues with multiple replicates	Disease cohort with >20+ samples	Plant or animal with multiple tissue types			
Target depth of coverage per sample	10 M reads per sample	5 M reads per sample	≤5 M reads per tissue (of same species)			
Comple multiple vin a 1	Sequel II/IIe system: Up to 2 samples per SMRT Cell 8M (2-plex)	Sequel II/IIe system: Up to 3 samples per SMRT Cell 8M (3-plex)	Sequel II/IIe system: Up to 3 tissue types per SMRT Cell 8M (3-plex)			
Sample multiplexing <sup>1</sup>	Revio system: Up to 4 samples per Revio SMRT Cell (4-plex)	Revio system: Up to 8 samples per Revio SMRT Cell (8-plex)	Revio system: Up to 8 tissue types per Revio SMRT Cell (8-plex) <sup>2</sup>			
Expected data throughput (per SMRT	Sequel II/IIe system: 15 M reads per SMRT Cell 8M divided by N samples					
Cell)	R	evio system: 40 M reads divided by N sample	es			
Kinnex library prep protocol	Procedure & checklist – Preparing Kinnex libraries using the Kinnex full-length RNA kit (103-238-700)					
Total RNA input into Kinnex library prep workflow	300 ng total RNA (RIN ≥7) for 1st-strand cDNA synthesis					
SMRT Link data analysis workflows	Read Segmentation and Iso-Seq analysis application with option to "pool reads and cluster together" to get a master isoform classification file with per-sample full-length read counts					
Community data analysis tools	Annotation & quantification: PIGEON, SQANTI3, Differential analysis: TappAS, Fusion calling: pbfusion, Visualization: SWAN					



<sup>&</sup>lt;sup>1</sup> Kinnex concatenation kit (103-071-800) can support up to 48-plex sample multiplexing through the combined use of 12 different barcoded cDNA amplification primers and 4 different barcoded Kinnex terminal SMRTbell adapters during Kinnex full-length RNA library construction.

<sup>&</sup>lt;sup>2</sup> If targeting <5 M transcripts reads per sample → can multiplex up to 12 tissues types per Revio SMRT Cell.

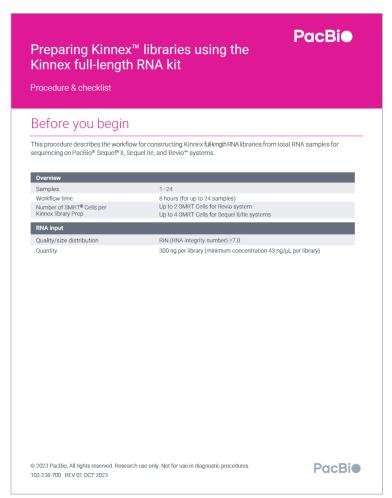
# Kinnex full-length RNA library preparation workflow details

# Procedure & checklist – Preparing Kinnex libraries using the Kinnex full-length RNA kit (103-238-700)

Procedure & checklist <u>103-238-700</u> describes the workflow for constructing Kinnex libraries from total RNA samples using the **Iso-Seq express 2.0 kit** and **Kinnex full-length RNA kit** for sequencing on PacBio Sequel II/IIe & Revio systems

#### **Procedure & checklist contents**

- 1. Total RNA input QC recommendations and general best practices for reagent & sample handling.
- 2. Enzymatic workflow steps for cDNA synthesis and amplification.
- 3. Enzymatic workflow steps for construction of 8-segment Kinnex arrays from amplified cDNA.
- 4. Enzymatic workflow steps for DNA damage repair & nuclease treatment of Kinnex libraries.
- 5. Workflow steps for final cleanup of Kinnex SMRTbell libraries using SMRTbell cleanup beads.



PacBio Documentation (103-238-700)



# Kinnex full-length RNA library construction workflow overview

Procedure & checklist – Preparing Kinnex libraries using the Kinnex full-length RNA kit (103-238-700)

	Cleanup	DNA QC	Walk-away ti	me*	
1 Input RNA quality control	N/A	Qubit RNA HS assay Bioanalyzer RNA 6000 Nano kit	N/A		
↓ O Safe stop			,	_	V
2 cDNA synthesis + cleanup	1.3X SMRTbell cleanup beads	N/A	1.5 hrs		
	cicariap scade				
3 cDNA amplification + cleanup	0.9X SMRTbell	Qubit dsDNA HS assay	1.5 hrs	Day 1	
↓ O Safe stop	cleanup beads	Bioanalyzer HS DNA assay		(~5 hrs*)	See Line
4 Kinnex PCR + cleanup	1.05X SMRTbell cleanup beads	Qubit dsDNA HS assay	1.5 – 2.0 hrs		
<u> </u>	Cleariup beaus			_	
5 Kinnex array formation + cleanup	1.0X SMRTbell cleanup beads	N/A	2.0 hrs	٦	
↓ <b>O</b> Safe stop	cicariap scade			Day 2	
6 Nuclease treatment + cleanup	1.0X SMRTbell cleanup beads	Qubit dsDNA HS assay Femto Pulse gDNA 165 kb kit	0.6 hrs	(~3 hrs*)	



# General best practices recommendations for preparing Kinnex full-length RNA libraries

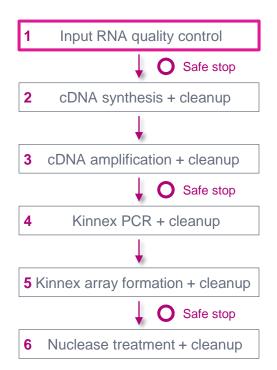
### Reagent and sample handling

- Take care to accurately pipette SMRTbell cleanup beads because small changes in volume can significantly alter the size distribution
  of your sample.
- Equilibrate the SMRTbell cleanup beads at room temperature for 30 mins prior to use.
- In cDNA amplification and Kinnex PCR, keep sample(s) on ice until thermal cycler lid has reached 105°C to avoid digestion of primers by polymerase exonuclease activity.
- This workflow takes ~8 hrs to complete.
  - If a stop is necessary, refer to the workflow for safe stopping points.



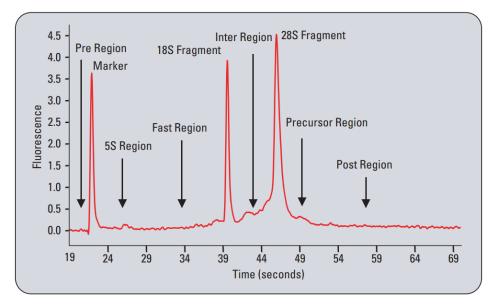
# Input RNA quality control

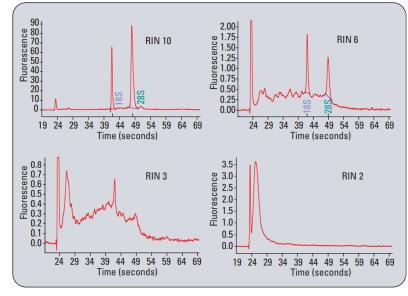
Input RNA quality control is highly recommended before proceeding to the MAS-Seq library prep workflow



**Left:** Bioanalyzer electropherogram detailing the regions that are indicative of RNA quality. **Right:** Sample electro-pherograms corresponding to different RNA Integrity Number (RIN) scores. Samples range from intact (RIN 10), to degraded (RIN 2). Images from Agilent Application Note: RNA Integrity Number (RIN) — Standardization of RNA Quality Control (5989-1165EN)

- 300 ng of total RNA per library (minimum concentration 43 ng/µL per library) is required for this procedure
- Sample QC of input total RNA samples should be assessed by measuring RNA Integrity Number (RIN) using a Bioanalyzer 2100 instrument (Agilent Technology) with RNA 6000 Nano kit<sup>1</sup>
  - RIN ≥7.0 (ideally ≥8.0) is sufficient for Kinnex full-length RNA protocol
  - Samples with RIN <7.0 can be processed, but risk of significant underperformance or even failure is greatly increased</li>
- RIN score (1 to 10) is related to ratio of the area under 28s and 18s fragment peaks and also takes into
  account signal intensity above baseline in the Inter-Region and Fast Region since this is where
  degradation products appear
- Higher RIN numbers are correlated with better overall sample quality and lower degradation

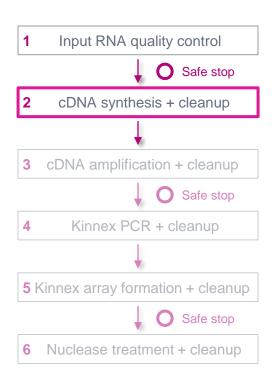


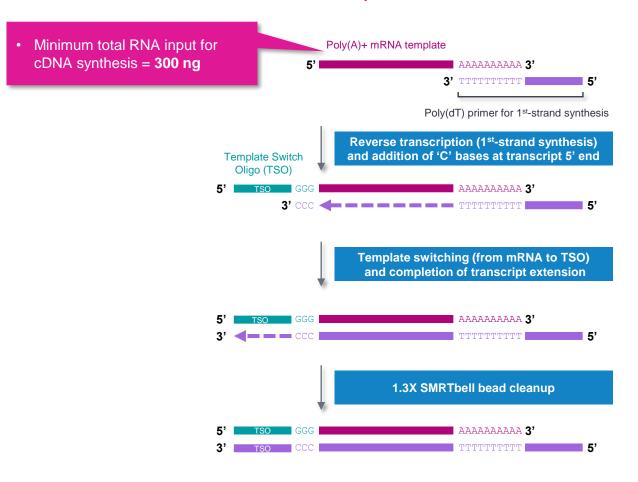




# cDNA synthesis + cleanup

In this step, total RNA samples are converted to first-strand cDNA products

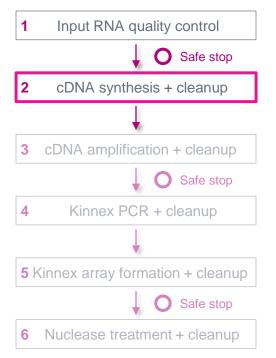




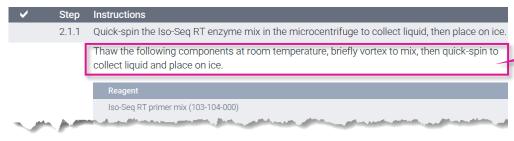


# cDNA synthesis + cleanup (cont.)

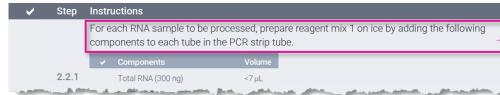
#### Procedural notes



#### 2.1 Thawing reagents for first-strand synthesis



#### 2.2 Primer annealing for first-strand synthesis



#### 2.3 Reverse transcription and template switching



#### 2.4 1.3X SMRTbell bead cleanup

✓ Step	Instructions
2.4.1	For each sample, add 29 $\mu$ L of elution buffer to the 21 $\mu$ L reverse transcription and template switching reaction (Section 2.3) for a total volume of 50 $\mu$ L.
2.4.2	Add 65 µL of resuspended, room-temperature SMRTbell cleanup beads.
2.4.3	Mix beads by pipetting 10 times or until evenly distributed.

 After thawing specified reagents on ice, perform a quick spin to collect liquid, then place on ice

 Set up primer annealing reaction mix on ice and then transfer to thermal cycler for incubation

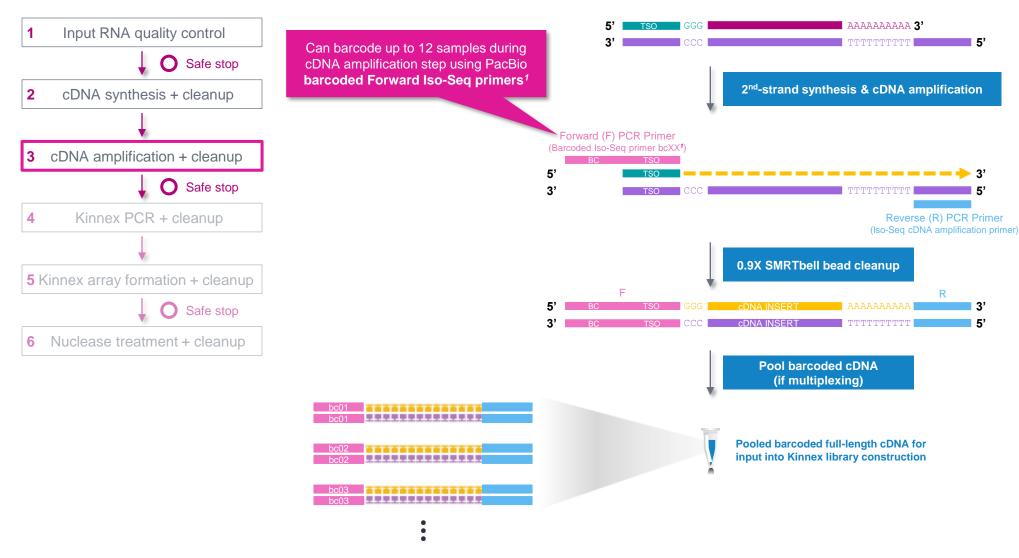
Set up reverse transcription reaction mix **on ice** and then transfer to thermal cycler for incubation

• Perform 1.3X SMRTbell bead cleanup



# cDNA amplification + cleanup

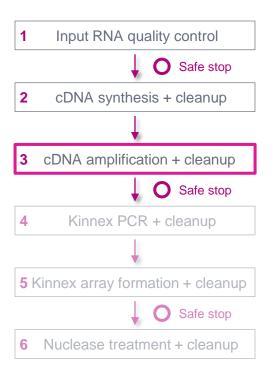
In this step, first-strand cDNA products are PCR-amplified and barcoded using barcoded Iso-Seq primers



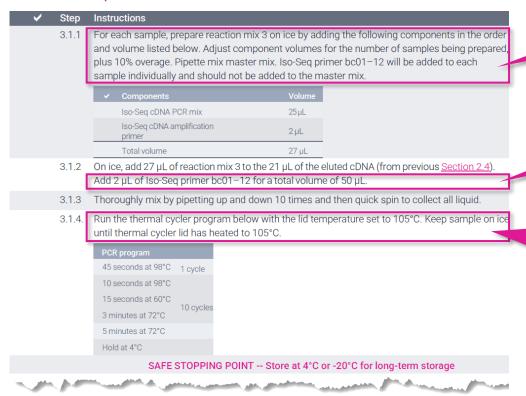


# cDNA amplification + cleanup

#### Procedural notes



#### 3.1 cDNA amplification



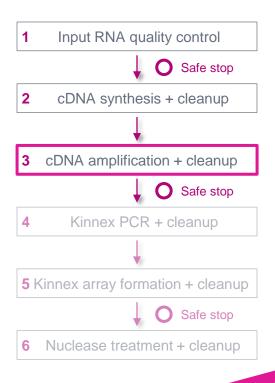
- Set up cDNA amplification reaction mix on ice
- Do not add Barcoded Forward Iso-Seq primers to master mix reaction
- Add desired Barcoded Forward Iso-Seq primer (select one of bc01-bc12) to each individual cDNA amplification reaction

 After setting up reactions on ice, add PCR reactions to thermal cycler after the lid has preheated to 105°C to avoid digestion of primers by polymerase exonuclease activity



# cDNA amplification + cleanup (cont.)

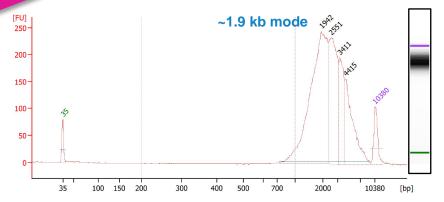
#### Procedural notes



3.2 Cleanup of amplified cDNA using 0.9X SMRTbell Cleanup beads

	<b>~</b>	Step	Instructions		
		3.2.1	Add 45 $\mu$ L (0.9x) of resuspended, room-temperature SMRTbell cleanup beads to the 50 $\mu$ L of cDNA amplified reaction from Section 3.1. The correct ratio of beads to sample is critical at this step.		
		3.2.2	Mix beads by pipetting 10 times or until evenly distributed.		
3.2.3 Quick-spin strip tubes in a microcentrifuge to collect liquid.					
		3.2.4	Leave at room temperature for 10 minutes to allow DNA to bind beads.		
	٦	4			
3.2.14 Slowly pipette off the cleared supernatant without disturbing the beads. Transfer 24 supernatant to a new strip tube. Discard the old strip tube with beads.					
			Recommended: Measure concentration and size distribution of each cDNA sample.		
		3.2.15	<ul> <li>Take a 1 µL aliquot from each strip tube. Dilute each aliquot with 4 µL of elution buffer.</li> <li>Measure DNA concentration with a Qubit fluorometer using the 1x dsDNA HS kit.</li> <li>Dilute 1:4 dilution further to 1.5 ng/µL based on the Qubit reading if needed.</li> <li>Run 1 µL on an Agilent Bioanalyzer using a High Sensitivity DNA kit.</li> </ul>		
		3.2.16	The expected recovery after cDNA amplification SMRTbell clean-up is >100 ng. A minimum of 55 ng of total cDNA is recommended to proceed with Kinnex PCR ( <u>Step 4</u> ). If less than 55 ng but more than 25 ng is recovered, proceed with Kinnex PCR but expect lower yields. Do not proceed with less than 25 ng.		
	ماول	1	The state of the s		

- Perform 0.9X SMRTbell bead cleanup<sup>1</sup>
- Perform DNA concentration QC using Qubit ds DNA HS assay and DNA sizing QC using Bioanalyzer



Example Bioanalyzer DNA sizing QC analysis results for amplified full-length cDNA generated from a universal human RNA reference (UHRR) total RNA sample.

 If multiplexing, perform equal-mass pooling of each barcoded cDNA sample after 0.9X SMRTbell bead cleanup

- Expected yield of purified cDNA product is >100 ng
- Minimum cDNA amount needed to proceed with Kinnex PCR = 55 ng (do not proceed with <25 ng)</li>

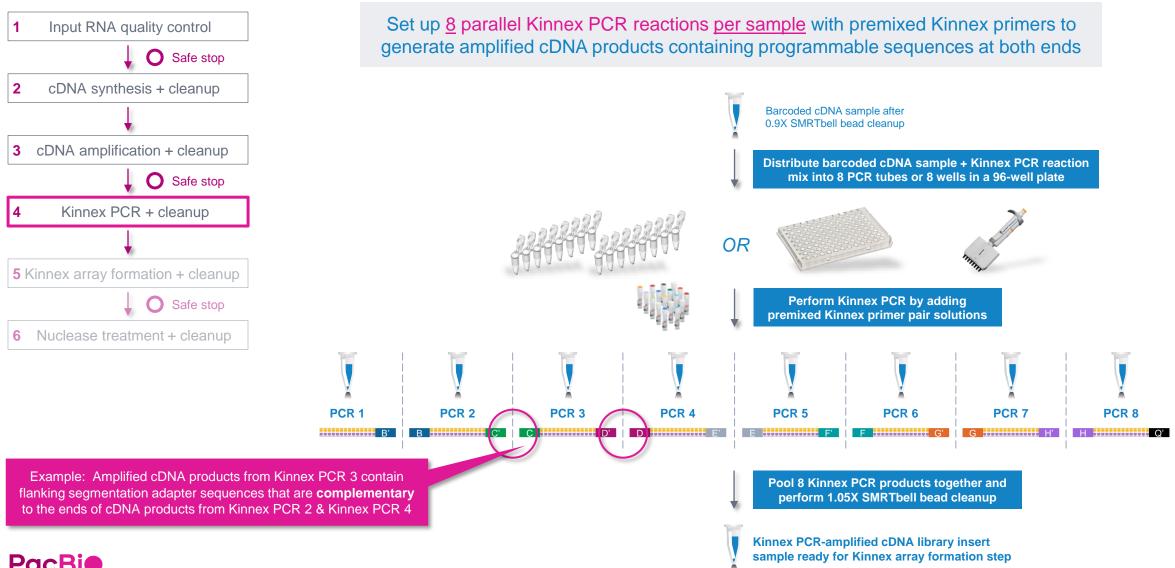
#### 3.3 Pooling barcoded cDNA (skip if not multiplexing)

✓	Step	Instructions
	3.3.1	Using the concentration reading from the Qubit fluorometer, pool an equal mass of each barcoded cDNA sample for a total mass of 55 ng. Store any remaining purified, amplified barcoded cDNA at 4°C for future use.
	3.3.2	Quick-spin the tube strip in a microcentrifuge to collect liquid.
	3.3.3	Proceed to next step of the protocol.



# **Kinnex PCR**

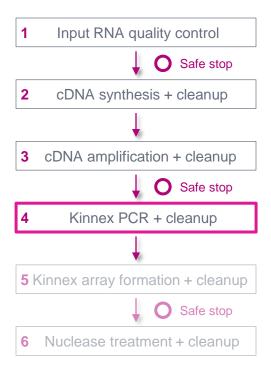
In this step, incorporate programmable Kinnex segmentation adapter sequences into amplified cDNA products



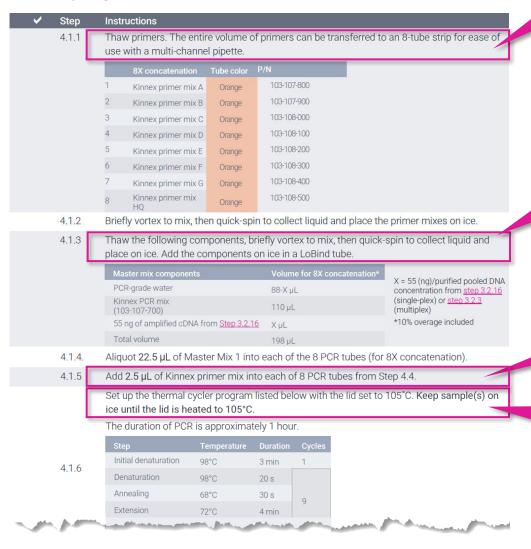


# **Kinnex PCR (cont.)**

#### Procedural notes







 Can transfer entire volume of primers to PCR tubes for ease of use with multichannel pipettes (8 primer mix tubes)



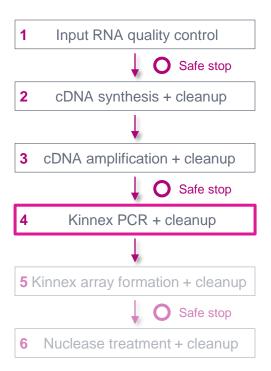
- Set up Kinnex PCR reactions ON ICE
- PCR polymerase 3'→5' exonuclease activity negatively impacts amplification yield if prepared at room temp.
- Critical step! Correct setup of all 8 Kinnex PCR reactions is required – any missing/incorrect MAS primer pairs will result in no/low SMRTbell yield

 Set up on ice and add PCR reaction to thermal cycler after lid has preheated to 105°C to avoid digestion of primers by polymerase exonuclease activity



# **Kinnex PCR (cont.)**

#### Procedural notes



#### 4.2 Pooling of 8 Kinnex PCR products and 1.05X SMRTbell cleanup

<b>✓</b>	Step	Instructions
	4.2.1	Add 23 $\mu$ L from each of the 8 PCR reactions into a 1.5 mL tube for a total volume of 184 $\mu$ L. Appendix a polymer of each PCR product is necessary for efficient array assembly.
	4.2.2	Add 193 $\mu$ L (1.05X v/v) of resuspended, room-temperature SMRTbell cleanup beads to a tube of pooled Kinnex PCR amplicon. The correct ratio of beads to pooled sample is critical at this step.
	4.2.3	Pipette-mix the beads until evenly distributed.
	4.2.4	Quick-spin the tube in a microcentrifuge to collect liquid.
	4.2.5	Leave at room temperature for 10 minutes to allow the DNA to bind beads
	4.2.6	Place the tube in a magnetic separation rack until the beads separate fully from the solution.
	4.2.7	Slowly pipette off the cleared supernatant without disturbing the beads. Discard the supernatant.
	4.2.8	Slowly dispense 200 $\mu$ L, or enough to cover the beads, of freshly prepared 80% ethanol into the tube. After 30 seconds, pipette off the 80% ethanol and discard.
	4.2.9	Repeat the previous step.
	4.2.10	Remove residual 80% ethanol:  Remove the tube from the magnetic separation rack.  Quick-spin the tube in a microcentrifuge.  Place the tube back in the magnetic separation rack until the beads separate fully from the solution.  Pipette off residual 80% ethanol and discard.
	4.2.11	Remove the tube from the magnetic rack. Immediately add 40 $\mu$ L of elution buffer to the tube and resuspend the beads by pipetting 10 times or until evenly distributed.
	4.2.12	Quick-spin the tube in a microcentrifuge to collect liquid.
	4.2.13	Leave at room temperature for 5 minutes to elute DNA.
	4.2.14	Place tube in a magnetic separation rack until beads separate fully from the solution.
	4.2.15	Slowly pipette off the cleared supernatant without disturbing the beads. Transfer supernatant to a <b>new LoBind tube</b> . Discard old tube with beads.
	4.2.16	Make a 1:10 dilution of the sample in elution buffer and measure DNA concentration with a Qubit fluorometer using the 1x dsDNA HS kit. Typical yield is 6–12 μg.

Marie Commence of the State of

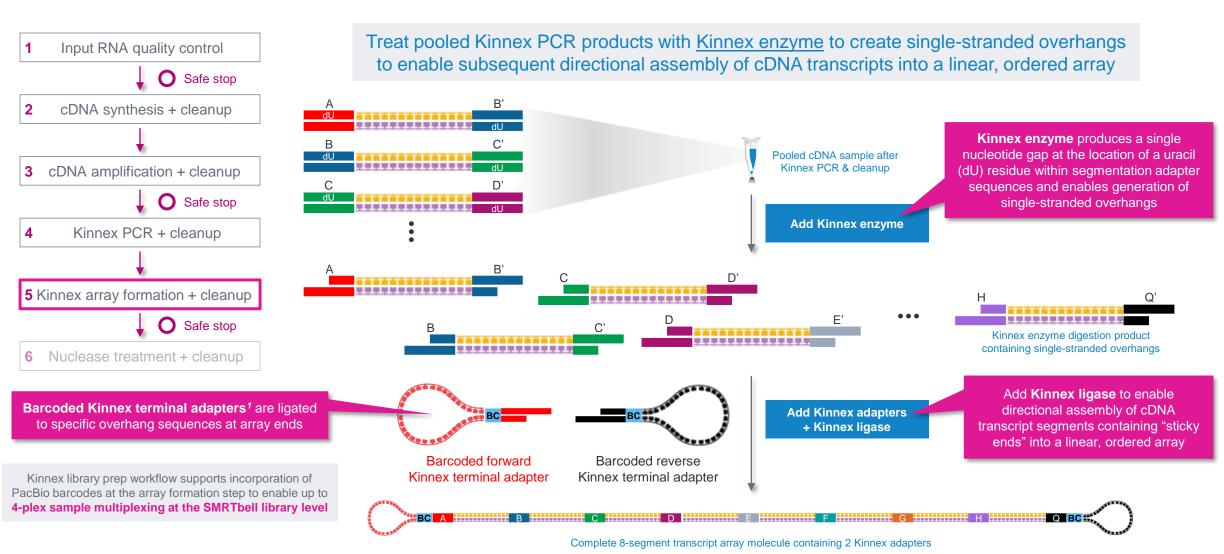
- Pool exactly 23 μL from each Kinnex PCR reaction in a clean 1.5 mL DNA LoBind tube<sup>1</sup> for a total combined volume of 184 μL
- Add exactly 193 μL of SMRTbell cleanup beads (1.05X)
- Kinnex PCR mix significantly increases stringency of SMRTbell clean up beads, so accurate pipetting is critical

 Perform DNA concentration QC to verify there is sufficient yield of Kinnex PCR products (min. 4 μg) to proceed to Kinnex array formation step



# **Kinnex array formation**

In this step, assemble cDNA transcripts ("segments") containing programmable ends into a linear array



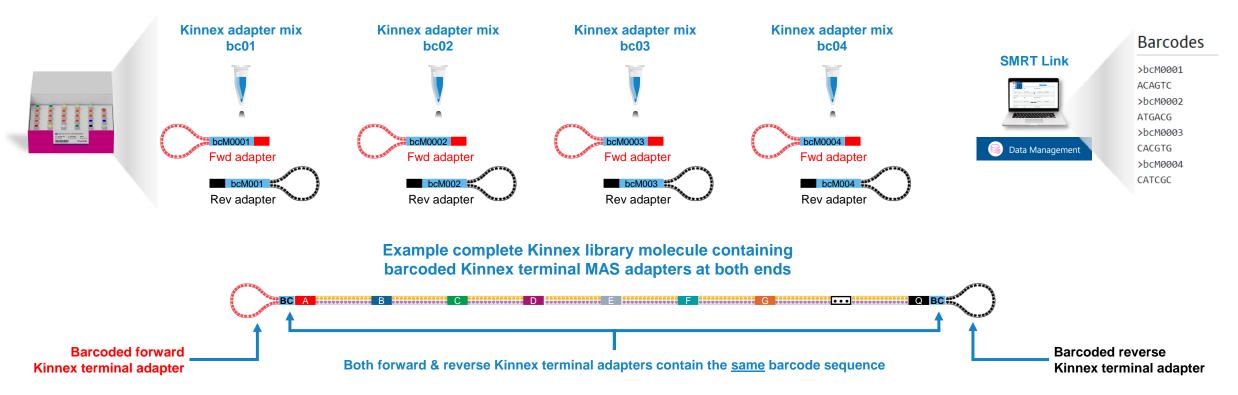


\* Note: Four barcoded terminal Kinnex adapters (Kinnex adapter bc01-04) are available for Kinnex array formation step. Note: Kinnex concatenation workflow is not compatible with standard SMRTbell adapters from SMRTbell prep kit 3.0 and is also not compatible with SMRTbell barcoded adapter plate 3.0.

# **Kinnex array formation (cont.)**

Kinnex terminal adapters incorporate barcode sequences to enable up to 4-plex sample multiplexing at the library level

- Kinnex adapters contain barcode sequences<sup>1</sup> to enable (optional) sample multiplexing at the SMRTbell library level (up to 4-plex)
  - Forward and reverse Kinnex adapter pairs are pre-mixed in Kinnex concatenation kits
  - Kinnex concatenation kits contain a total of 4 barcoded Kinnex adapter mixes (bc01-bc04) to enable multiplexing of up to 4 samples per SMRT Cell

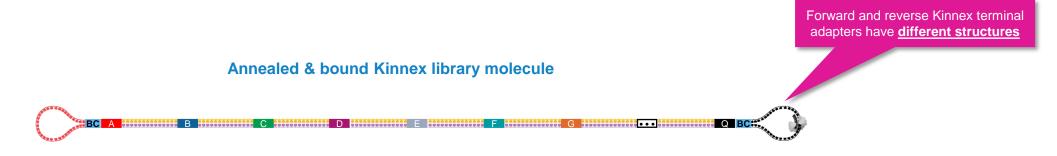




# **Kinnex array formation (cont.)**

Kinnex terminal adapters use a new design that enables improved SMRT sequencing performance

- Kinnex adapters enable:
  - Longer polymerase read length → Improved HiFi conversion rate (HiFi reads/Total *P1* reads)
  - Improved P1 loading efficiency



New Kinnex adapter design requires a different sequencing primer (Kinnex sequencing primer 103-179-000)



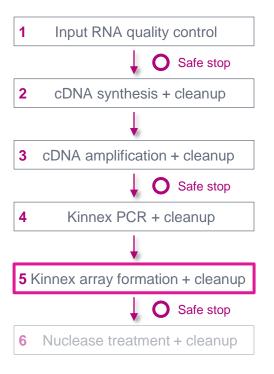


<sup>1</sup> If needed, Kinnex sequencing primer may be provided free-of-charge to any customers that are still using older Sequel II binding kits or older Revio polymerase kits that do not already contain the primer.

<sup>&</sup>lt;sup>2</sup> Revio polymerase kit includes SMRTbell cleanup beads.

# **Kinnex array formation (cont.)**

#### Procedural notes



#### 5.1 Kinnex array formation

<b>✓</b>	Step	Instructio	ns					
	5.1.1		nL PCR tube, add 4–8 µg of sample from <u>Step 4.2.15</u> , in 39 µL of volume (102–205 ilute with elution buffer going into this step if the sample is too concentrated.					
						ne barcode per library pre		
		Note: if combining multiple barcoded Kinnex libraries for sequencing, make sure each librar						
				erent Kinnex l		•		
	5.1.2	Add the fo	ollowing con	mponents in th	the listed ord	der.		
		If process	sing multiple	samples, ma	ake a maste	r mix with 10% overage. F	Pipette mix master mix.	
			Component	s		Volume		
			Kinnex array 110-300)	and repair buffe	er (103-	7.0 µL		
			Kinnex enzyr	me (103-110-400	10)	4.0 µL		
			Kinnex ligas	e (103-110-500)		6.0 μL		
			Total RM1 v	olume		17 μL		
-	1	-	سیں کہ	the same of	المستول عوال	and the same of the	A CONTRACTOR OF THE PARTY OF TH	
-	-		-					
	5.1.4	After running the Kinnex primer digestion/ligation program, add 2 $\mu$ L of DNA repair mix directly to the Kinnex primer digestion/ligation sample.						
	5.1.5	Thoroughly mix by pipetting up and down 10 times and then quick spin to collect all liquid.						
	Run the DNA Damage Repair Program with the lid set to >55°C.							
	5.1.6	Step 7	Temperature	Duration				
	3.1.0	1	45°C	30 min				

- Recommended minimum input requirement to proceed with Kinnex array formation is **4 μg** of Kinnex PCR amplicons (from Step 4)
  - Proceeding with <4  $\mu g$  is **not recommended** since lower input amounts may lead to insufficient final library yields to enable optimal sequencing results

- **IMPORTANT:** If combining multiple barcoded Kinnex libraries for sequencing, make sure each library uses one of the 4 different Kinnex barcoded adapters
- Perform DNA Damage Repair step to repair nicked / damaged DNA sites within newly formed Kinnex array products

• Perform **1X** SMRTbell bead cleanup at room temp.

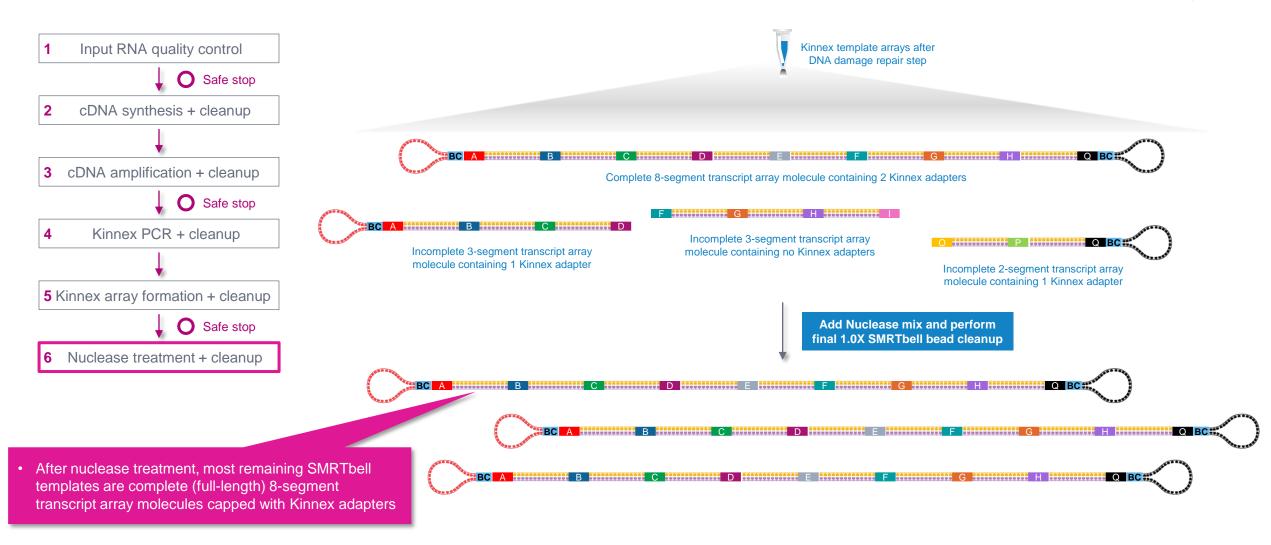
#### 5.2 1X SMRTbell bead cleanup

✓	Step	Instructions
	5.2.1	Add 1X v/v (60 $\mu$ L) of resuspended, room temperature SMRTbell cleanup beads to each sample.
	5.2.2	Pipette-mix the beads until evenly distributed and quick-spin in a microcentrifuge to collect liquid.
-	1	the control of the co



# **Nuclease treatment & cleanup**

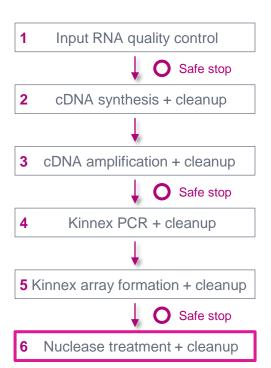
Perform nuclease treatment and final SMRTbell bead cleanup to remove incomplete SMRTbell template arrays





# **Nuclease treatment & cleanup (cont.)**

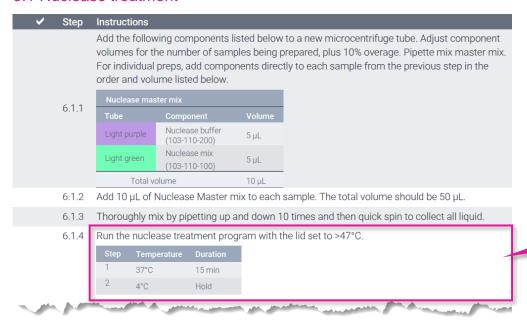
#### Procedural notes



 Perform final 1X SMRTbell bead cleanup at room temp.

Final Kinnex library yield is typically sufficient to load ≥2 SMRT Cells

#### 6.1 Nuclease treatment



#### 6.2 1X SMRTbell bead cleanup

<b>~</b>	Step	Instructions
	6.2.1	Add 50 µL SMRTbell cleanup beads to each sample from the previous step. Pipette-mix the beads until evenly distributed.
	6.2.2	Quick-spin the tube strip in a microcentrifuge to collect all liquid.
1	**	
	6.2.14	Take a 1 µL aliquot from each tube. Make a 1:5 dilution of the sample in elution buffer and measure DNA concentration with a Qubit fluorometer using the 1x dsDNA HS kit. Calculate the total mass. Expect 10-25% recovery of the starting Kinnex-PCR product.
		Recommended: Further dilute each aliquot to 250 pg/µL with the Femto Pulse dilution buffer.  Measure the final SMRTbell library size distribution with a Femto Pulse system.

And the second particular property of the second

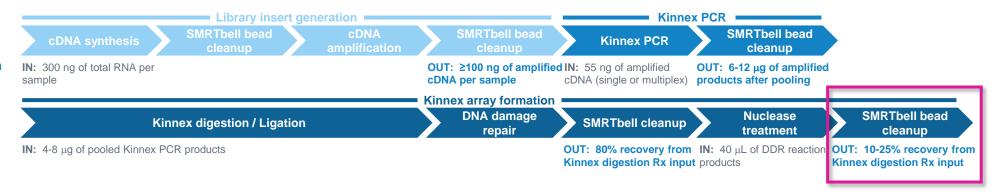
- Perform nuclease treatment for 15 min
- Perform **DNA concentration QC** on final purified Kinnex library using Qubit dsDNA HS assay
  - Typical final SMRTbell library yield from 4 8 μg of input DNA into Kinnex array formation is ~10 – 25% – a much higher observed yield might suggest incomplete digestion of partial SMRTbell templates
  - Troubleshooting tip: If SMRTbell library yield is higher than expected and P1 loading is lower than expected, consider repeating the nuclease treatment step
- Perform DNA sizing QC on final purified Kinnex library using a Femto Pulse system (expected final library insert size is ~12 – 16 kb)

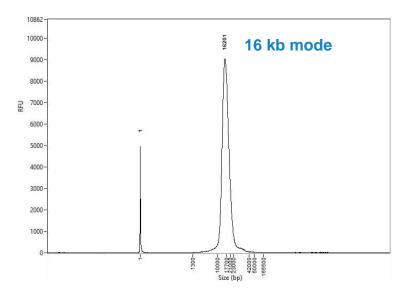


# Kinnex full-length RNA library prep inputs & expected step yields

## Final Kinnex library yield is typically sufficient to load ≥2 SMRT Cells

Preparing Kinnex libraries using the Kinnex full-length RNA kit (103-238-700)





Example Femto pulse DNA sizing QC analysis results for final Kinnex full-length RNA library prepared with human universal human reference RNA (UHRR) total RNA sample.

#### **Example Kinnex full-length RNA library prep yields**

Total RNA input for cDNA synthesis	300 ng
cDNA input for Kinnex array formation	5900 ng
Post-nuclease treatment & final library cleanup yield (%)	1460 ng (24.7%)

<sup>&</sup>lt;sup>1</sup> Post-nuclease treatment & final cleanup yields typically ranged from ~10% to ~25% when using UHRR total RNA samples for Kinnex full-length RNA library construction.

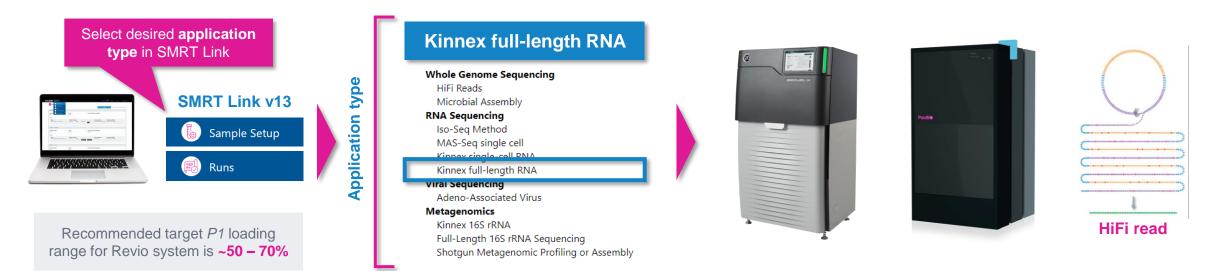
Final Kinnex library yield is typically sufficient to load ≥2 SMRT Cells





# Sample Setup & Run Design recommendations for Kinnex full-length RNA libraries

SMRT Link v13 supports Kinnex full-length RNA sequencing preparation & analysis workflow for PacBio systems

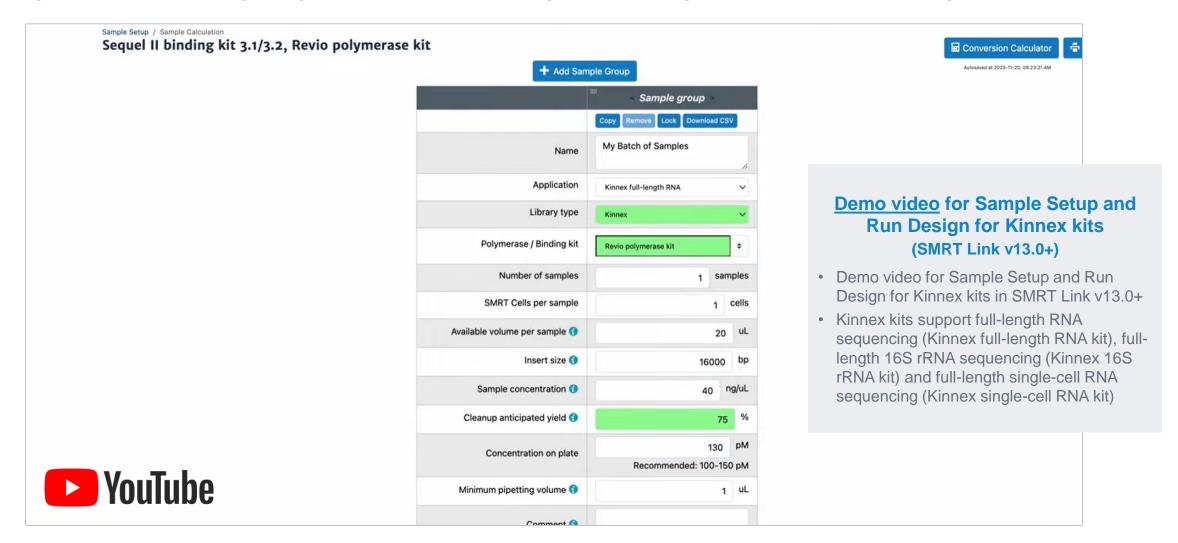


SMRT Link module	Key setup parameters For Kinnex libraries	Sequel II/IIe system recommended settings for Kinnex libraries	Revio system recommended settings for Kinnex libraries
Sample setup	Library type	Kinnex	
	Primer	Kinnex sequencing primer	
	Binding/Polymerase kit <sup>1</sup>	Sequel II binding kit 3.2 (includes Kinnex sequencing primer)	Revio polymerase kit (includes Kinnex sequencing primer)
	Concentration on plate	40 – 60 pM	100 – 150 pM
Runs → Run design	Adapter / Library type	SMRTbell Adapter Design = SMRTbell Kinnex Prep Kit	Library type = Kinnex
	Movie collection time	30 hrs	24 hrs
	Use adaptive loading	YES	
	On-instrument CCS	CCS Analysis Output - Include Low Quality Reads = NO CCS Analysis Output - Include Kinetics Information = NO	Consensus Mode = MOLECULE

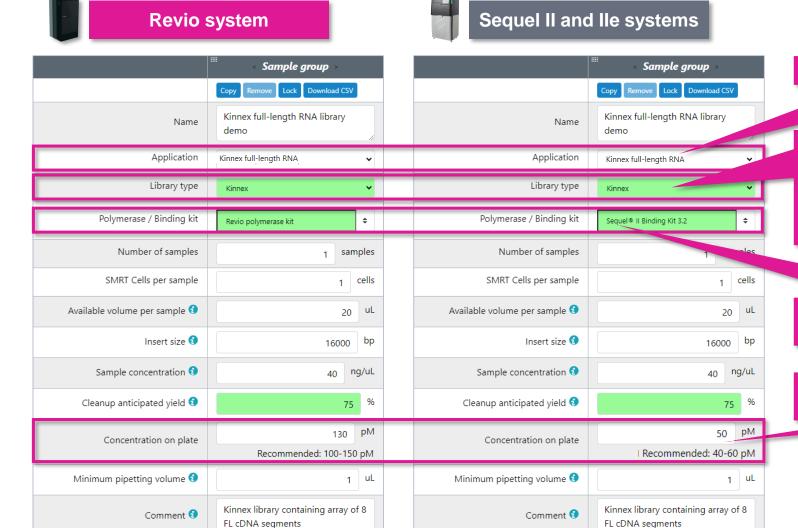


# SMRT Link Sample Setup and Run Design for Kinnex kits video demonstration

Video demonstration of SMRT Link Sample Setup and Run Design setup procedure for Kinnex kits supporting full-length RNA sequencing, single-cell RNA sequencing and full-length 16S rRNA sequencing



# SMRT Link Sample Setup procedure for Kinnex full-length RNA libraries



Select application type to autofill fields in green

#### **IMPORTANT:** Specify **Library type = Kinnex**

- Library type field determines sequencing primer type to use for annealing step
  - → Kinnex libraries require use of **Kinnex sequencing** primer<sup>1</sup>

Select Revio polymerase kit for Revio system and Sequel II Binding Kit 3.2 for Sequel II/IIe systems

Recommended starting OPLC range is 100 - 150 pM for Revio system and 40 - 60 pM for Seguel II/IIe systems

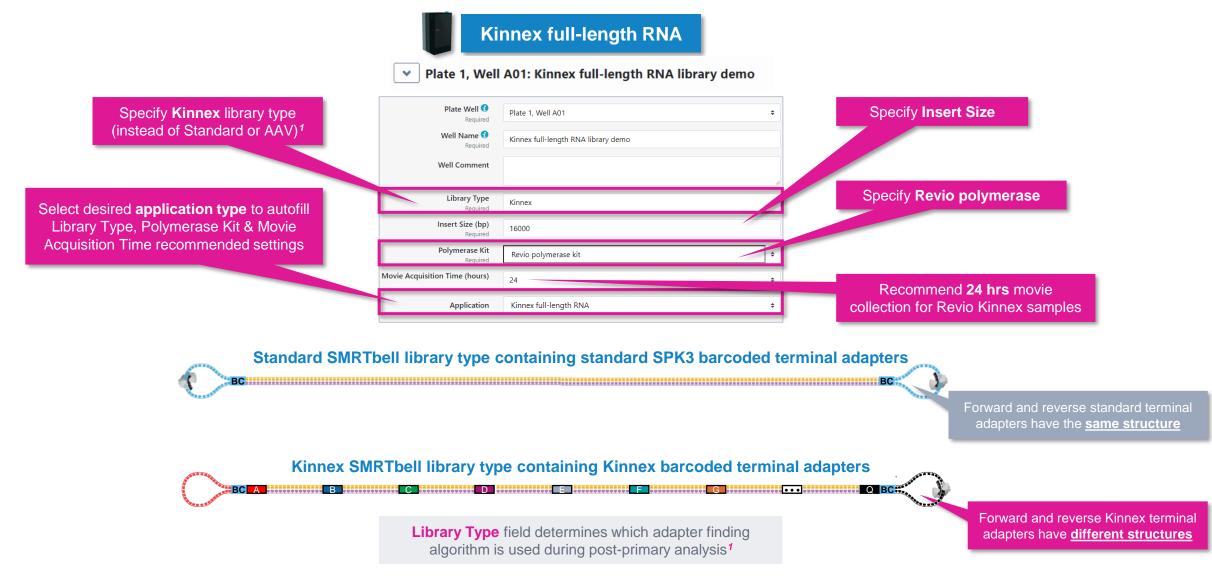
- Recommended target P1 loading range
- Revio system: ~50 70%
- Seguel II and IIe systems: ~60 80%



<sup>1</sup> Sample Setup Library type field specifies structure of SMRTbell library and determines sequencing primer type to use for annealing step. For Kinnex libraries, the forward and reverse Kinnex terminal SMRTbell adapters have different structures and require use of Kinnex sequencing primer for primer annealing step. (View Video tutorial – SMRT Link Sample Setup and Run Design setup procedure 24 for Kinnex kits at: YouTube Link)

# SMRT Link v13.0 Run Design procedure for Revio system

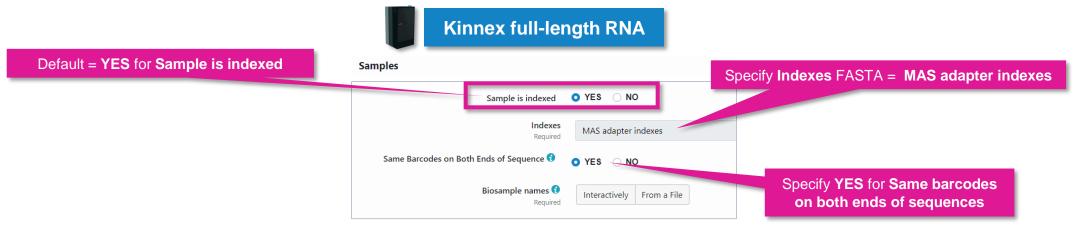
# Sample and run information



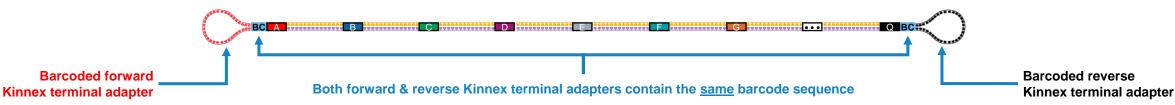
# SMRT Link v13.0 Run Design procedure for Revio system (cont.)

Sample indexing (barcoding) information





Example complete Kinnex library molecule containing barcoded Kinnex terminal MAS adapters<sup>1</sup> at both ends



#### **Example interactive biosample name specification for a multiplexed Kinnex library sample**

#### Barcode Selector and Sample Name Editor 3 Barcode Selector and Sample Name Editor 3 Available Barcodes Included Barcodes Available Barcodes Included Barcodes Filter... Filter... Filter... Barcode ↓î Bio Sample ↓↑ Barcode ↓↑ Barcode ↓↑ bcM0003--bcM0003 bcM0001--bcM0001 bcM0004--bcM0004 bcM0002--bcM0002



Filter...

Kinnex adapter-barcoded library 1

Kinnex adapter-barcoded library 2

Kinnex
MAS adapter indexes

>bcM0001
ACAGTC
>bcM0002
ATGACG
>bcM0003
CACGTG
>bcM0004
CATCGC

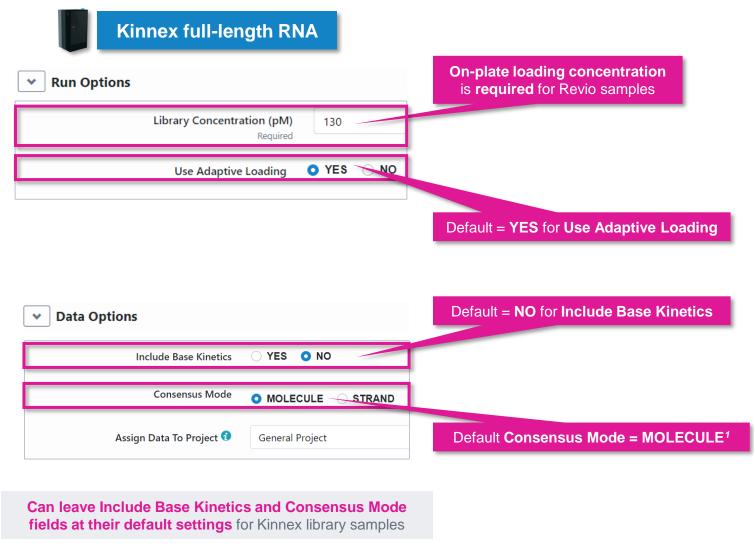


<sup>1</sup> Four barcoded terminal Kinnex adapters (Kinnex adapter bc01-04) are available for Kinnex array formation step. Kinnex adapter barcode sequences can be downloaded from <u>SMRT Link</u> Data Management module.

### SMRT Link v13.0 Run Design procedure for Revio system (cont.)

Run options and data options







### SMRT Link Run Design procedure for Sequel II/IIe systems

Sample information and run information





- Select desired Kinnex application from the Application field drop-down menu
- The following fields are auto-populated with default recommended values and high-lighted in green:
  - SMRTbell Adapter Design

    → SMRTbell Kinnex Prep Kit

    □ Binding Kit

    → Sequel II Binding Kit 3.2

    □ Sequencing Kit

    → Sequel II Sequencing Plate 2.0 (4 rxn or 1 rxn)

    □ DNA Control Complex

    → Sequel II DNA Internal Control Complex 3.2

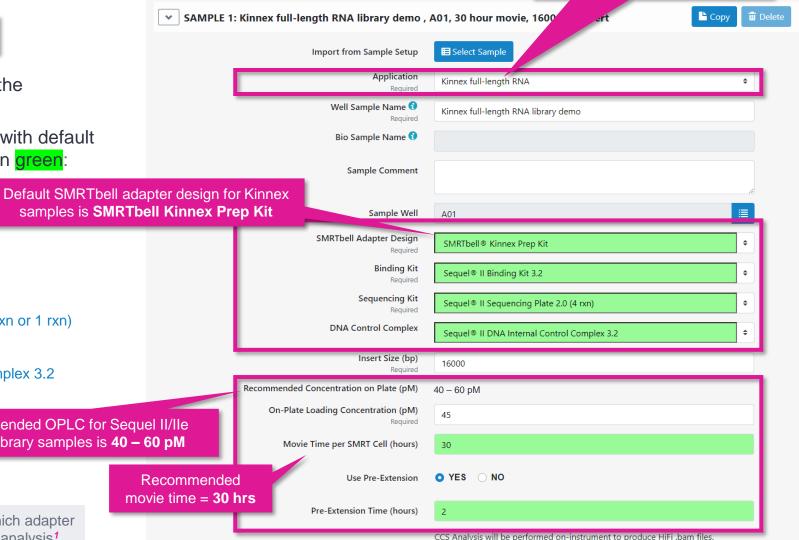
    □ Movie Time per SMRT Cell

    → 30 hrs

    □ Pre-Extension Time

    → 2 hrs

**SMRTbell Adapter Design** field determines which adapter finding algorithm is used during post-primary analysis<sup>1</sup>



Example sample information entered into a Sequel IIe system run design worksheet for a Kinnex full-length RNA library sample.



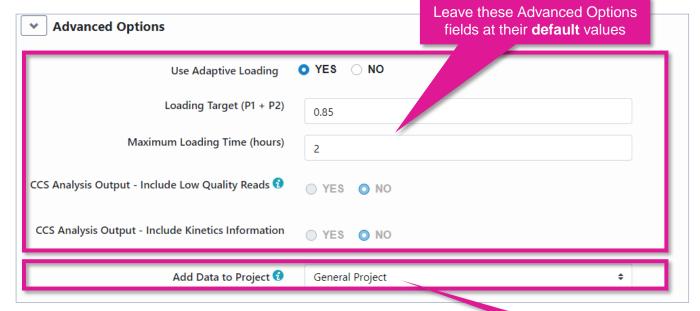
### SMRT Link Run Design procedure for Sequel II/IIe systems (cont.)

### Advanced options



### Kinnex full-length RNA

- For all Kinnex library samples, leave the following Advanced Options fields at their default settings
  - Use Adaptive Loading
    - $\rightarrow \ \ \mathsf{YES}$
  - Loading Target (P1 + P2)
    - $\rightarrow$  0.85
  - Maximum Loading Time
    - $\rightarrow$  2 hours
  - □ CCS Analysis Output Include Low Quality Reads
    - $\rightarrow$  NO
  - □ CCS Analysis Output Include Kinetics Information
    - $\rightarrow$  NO
  - Pre-Extension Time
    - $\rightarrow$  2 hrs
- If desired, specify to use an alternative project folder for the Add Data to Project field

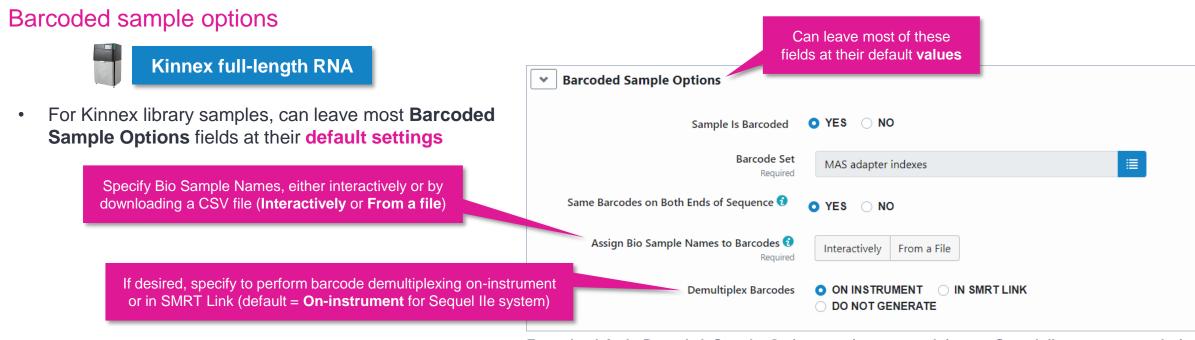


Example default Advanced Options settings entered into a Sequel IIe system up design worksheet for a Kinnex full-length RNA library sample.

Can specify to use a different Project folder



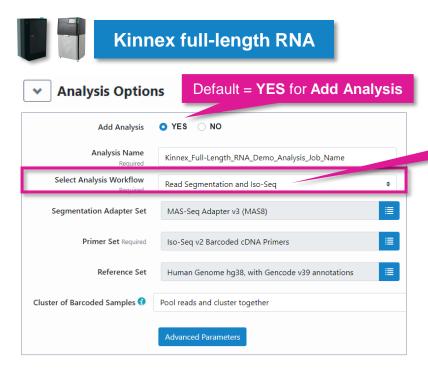
### SMRT Link Run Design procedure for Sequel II/IIe systems (cont.)



Example default Barcoded Sample Options settings entered into a Sequel IIe system run design worksheet for a Kinnex full-length RNA library sample.

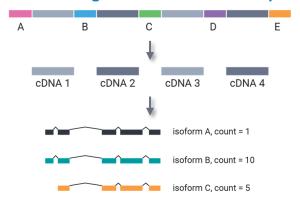


### SMRT Link v13.0 Run Design analysis options for Revio system and Sequel II/IIe systems



**Analysis Workflow** is automatically filled in (Default = Read Segmentation and Iso-Seq)

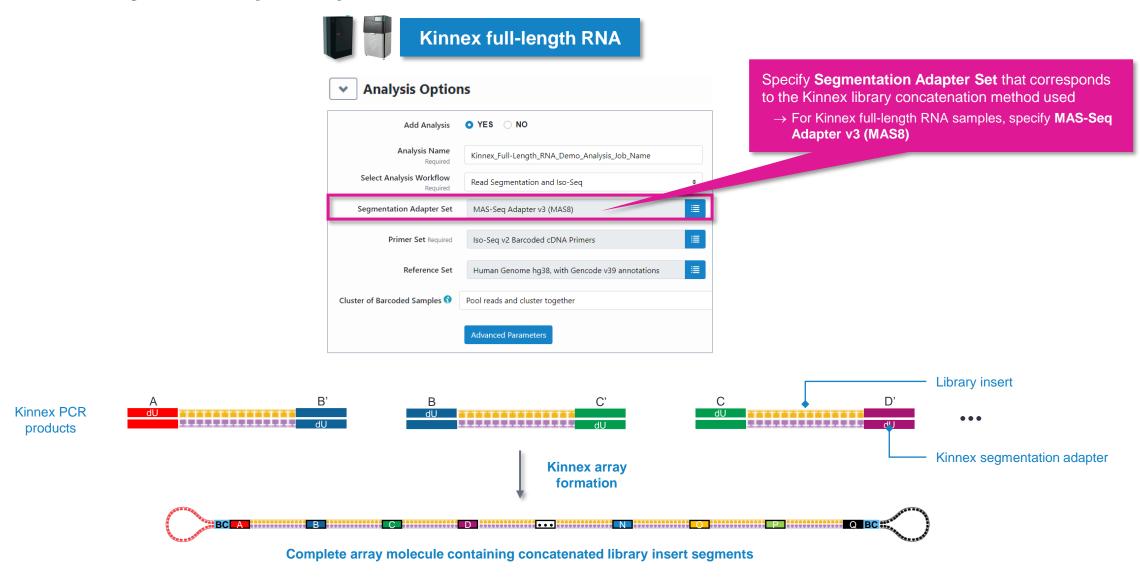
#### **Read Segmentation and Iso-Seq**



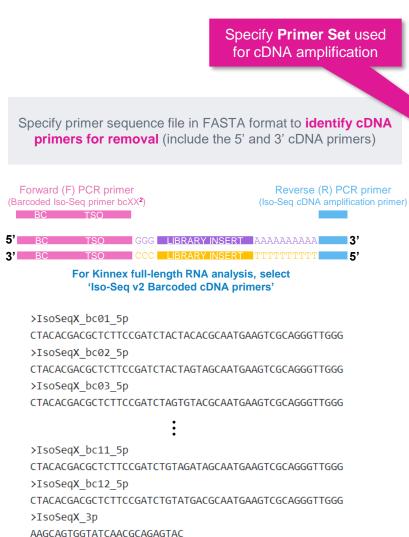
Perform isoform-classification analysis to identify novel genes & isoforms with abundance information (raw counts & normalized counts per million)

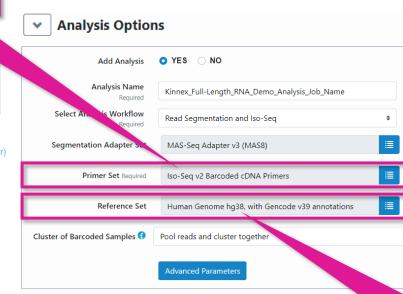


### SMRT Link v13.0 Run Design analysis options for Revio system and Sequel II/IIe systems (cont.)



### SMRT Link v13.0 Run Design analysis options for Revio system and Sequel II/IIe systems (cont.)





Kinnex full-length RNA

Specify reference genome & annotation sets to align high quality isoforms to, and to collapse isoforms mapped to the same genomic loci.



#### Specify **Reference** Set. Default sets are:

- Human Genome hg38, with Gencode v39 annotations
- Mouse Genome mm39, with Gencode vM28 annotations



SMRT Link v13.0 Run Design analysis options for Revio system and

Sequel II/IIe systems (cont.)

Specify how to perform read clustering for barcoded samples

Kinnex full-length RNA **Analysis Options** O YES O NO Add Analysis **Analysis Name** Kinnex\_Full-Length\_RNA\_Demo\_Analysis\_Job\_Name Select Analysis Workflow Read Segmentation and Iso-Seq MAS-Seq Adapter v3 (MAS8) Segmentation Adapter Set Iso-Seg v2 Barcoded cDNA Primers Primer Set Required Reference Set Human Genome hq38, with Gencode v39 annotations Cluster of Barcoded Samples @ Pool reads and cluster together Pool reads and cluster together -- select --Cluster reads separately Pool reads and cluster together Specification of Cluster of Barcoded Samples setting determines whether all FLNC reads will be pooled for **clustering** (Does not apply to non-barcoded samples.)

**Clustering options for Kinnex full-length RNA samples** Kinnex barcoded adapter demultiplexing (up to 4-plex) **Read segmentation** (Up to 8-fold Kinnex de-concatenation) Full-length cDNA barcode demultiplexing (up to 12-plex) Sample 1 ■ Sample 2 **Analyze pooled samples (default)** Analyze by sample classification.txt

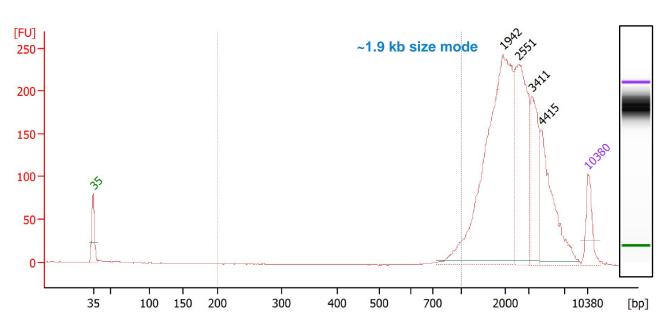
classification.txt



### **Example Kinnex full-length RNA library preparation QC results**

### Kinnex full-length RNA library prepared with human UHRR total RNA

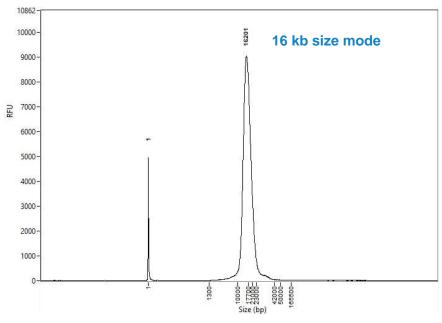
### Amplified full-length cDNA QC



Example Bioanalyzer DNA sizing QC analysis results for amplified full-length cDNA generated from a universal human RNA reference (UHRR) total RNA sample.

Final Kinnex library yield is typically sufficient to load ≥2 SMRT Cells

### Final Kinnex full-length RNA library QC



**Example Femto Pulse DNA sizing QC analysis results for final Kinnex full-length RNA library.** 

Total RNA input for cDNA synthesis	300 ng
cDNA input for Kinnex array formation	5900 ng
Post-nuclease treatment & final library cleanup yield (%)	1460 ng (24.7%)

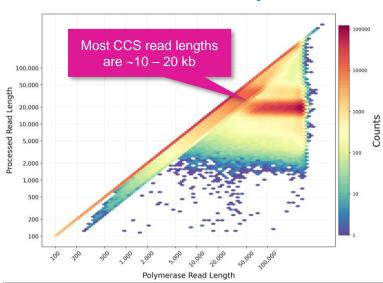
<sup>&</sup>lt;sup>1</sup> Post-nuclease treatment & final cleanup yields typically ranged from ~10% to ~25% when using UHRR total RNA samples for Kinnex full-length RNA library construction.



### Example sequencing performance for Kinnex full-length RNA libraries prepared with human cDNA

Sequel IIe system example data<sup>1</sup>

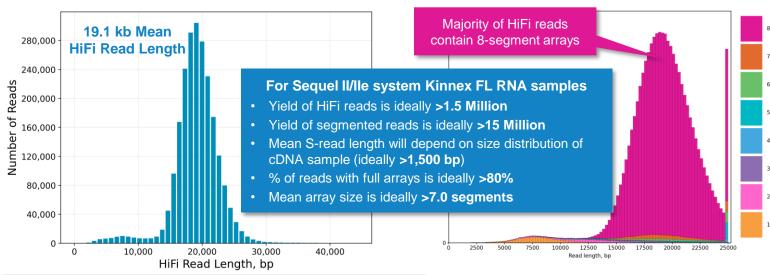
### **Raw Data Report**



Raw Base Yield	652 Gb
Mean Polymerase Read Length	115.36 kb
P0	28%
P1	71%
P2	1%

Example sequencing metrics for a Universal Human Reference RNA (UHRR) Kinnex full-length RNA library sample run on a Seguel IIe system with Binding Kit 3.2 (Polymerase 2.2) / 80 pM on-plate loading concentration (OPLC) / 30-hrs movie time / 2-hrs pre-extension time.

### HiFi Read Length



HiFi Reads	2.3 M
HiFi Base Yield	43.2 Gb
Mean HiFi Read Length	19.1 kb
Median HiFi Read Quality	Q32
HiFi Read Mean # of Passes	10

For UHRR Kinnex full-length RNA libraries, per-SMRT Cell 8M HiFi read counts typically ranged from ~2 - 3 Million depending on the final library insert size.

Reads	2,260,039
Segmented reads (S-reads)	17,213,165
Mean length of S-reads	2,420 bp
Percent of reads with full arrays	91.07%
Mean array size (concentration factor)	7.62

**Read Segmentation Metrics** 

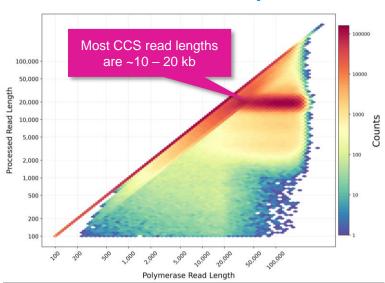
For UHRR Kinnex libraries, per-SMRT Cell 8M segmentation read counts were typically ~15 - 20 Million.



### Example sequencing performance for Kinnex full-length RNA libraries prepared with human cDNA

Revio system example data<sup>1</sup>

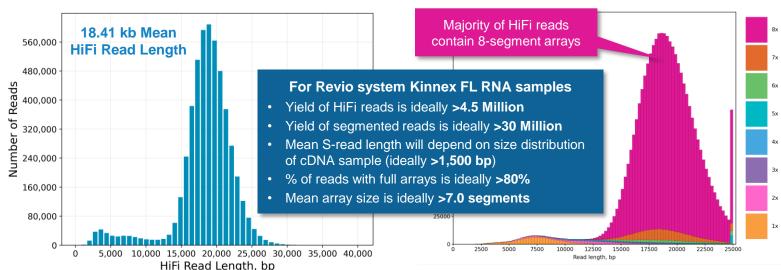
### **Raw Data Report**



Raw Base Yield	1,168 Gb
Mean Polymerase Read Length	58.6 kb
P0	16%
P1	79%
P2	5%

Example sequencing metrics for a Universal Human Reference RNA (UHRR) Kinnex full-length RNA library sample run on a Revio system with Revio polymerase kit / 130 pM on-plate loading concentration (OPLC) / 24-hrs movie time.

### HiFi Read Length



HiFi Reads	5.1 M
HiFi Base Yield	93.47 Gb
Mean HiFi Read Length	18.41 kb
Median HiFi Read Quality	Q28
HiFi Read Mean # of Passes	7

For UHRR Kinnex full-length RNA libraries, per-Revio SMRT Cell HiFi read counts were typically ~5 – 6 Million depending on the final library insert size and *P1* loading performance.

Input HiFi Reads	5,027,154
Segmented reads (S-reads)	37,216,151
Mean length of S-reads	2,393 bp
Percent of reads with full arrays	85.84%
Mean array size (concentration factor)	7.40

**Read Segmentation Metrics** 

For UHRR Kinnex libraries, per-Revio SMRT Cell segmentation read counts were typically  $\sim 30-45$  Million.





### Kinnex full-length RNA bioinformatics workflow overview

SMRT Link Read segmentation and Iso-Seq workflow processes HiFi reads generated from Kinnex full-length RNA libraries to produce classified isoforms with read counts that are compatible with tertiary analysis tools

### SMRT Link read segmentation and Iso-Seg workflow<sup>1</sup> HiFi reads S-reads TSO Transcript polyA **SMRT Link** Primer extraction Isoform clustering Mapping Known isoform A **Transcript** Known isoform B classification Novel isoform C

#### **Read segmentation**

 HiFi reads are segmented into individual segmented reads (S-reads) that represent the original cDNA sequences

#### **Primer extraction**

Primers and polyA tails are removed, but also used to orient the read into 5'
 → 3' orientation

#### **Isoform clustering**

- FLNC reads are clustered by their sequencing similarity to produce isoform consensus sequences
- This step is the last step of Iso-Seq analysis if no genome is provided

#### **Mapping**

 If a genome is provided, isoform consensus sequences from the previous step are mapped and further collapsed by their exonic structures to produce isoforms as GFF files for visualization

#### Transcript classification<sup>2</sup>

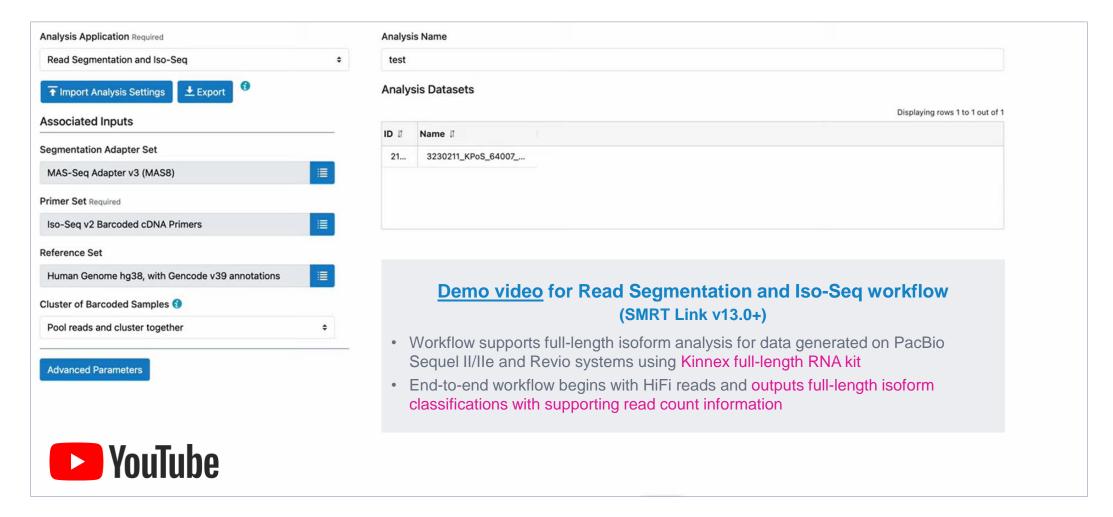
- If an annotation (e.g., Gencode) is provided, isoforms are classified against it using pigeon (the PacBio implementation of SQANTI3) to identify known and novel genes/isoforms
- The Iso-Seq workflow can jointly analyze pooled sample reads to produce a unified isoform annotation with per-sample read counts, both raw and normalized as counts per million (CPM)



<sup>1</sup> See SMRT Link User Guide (103-008-800) for detailed descriptions of parameter settings for Read Segmentation and Iso-Seq analysis application. (A video tutorial is also available for viewing.)

### SMRT Link Read Segmentation and Iso-Seq analysis video demonstration

Video demonstration of SMRT Link Read Segmentation and Iso-Seq application workflow for analysis of Kinnex full-length RNA samples





### Kinnex full-length RNA bioinformatics workflow recommendations

SMRT Link Read Segmentation and Iso-Seq workflow common considerations and recommendations for analysis of Kinnex full-length RNA data<sup>1</sup>

### Analysis recommendations for Iso-Seq data based on reference genome and annotation availability<sup>1</sup>

- With SMRT Link v13, the **Read segmentation and Iso-Seq workflow** analysis application supports human and mouse reference genomes and annotations to produce classified isoforms with read counts.
- If working with other organisms, see table below for analysis recommendations

Available reference or annotation	Analysis workflow recommendation
Human or mouse	<ul> <li>Use the Iso-Seq workflow with preloaded human / mouse annotation to get mapped, unique isoforms with classifications and read count information (FASTA, GFF, TXT).</li> </ul>
Model organism with good annotation	<ul> <li>Run Iso-Seq workflow with uploaded reference genome to get mapped, unique isoforms (FASTA, GFF)</li> <li>Generate pigeon-compliant annotation and use the command line for isoform classification with read count information (TXT)</li> </ul>
Non-model organism with genome	<ul> <li>Run Iso-Seq workflow with uploaded reference genome to get mapped, unique isoforms (FASTA, GFF)</li> </ul>
No genome	Run Iso-Seq workflow without reference genome to get unique isoforms (FASTA)



### Kinnex full-length RNA bioinformatics workflow recommendations (cont.)

SMRT Link Read Segmentation and Iso-Seq workflow common considerations and recommendations for analysis of Kinnex full-length RNA data<sup>1</sup>

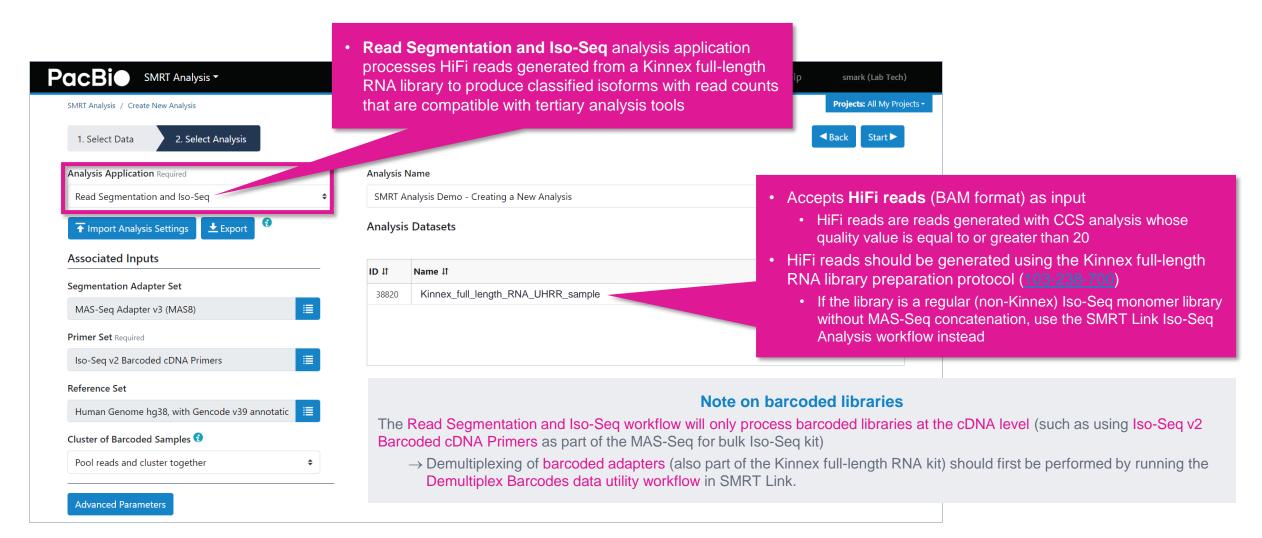
### Sequencing depth recommendations for Iso-Seq data based on experimental goals and study design

Example application	Human genetics disease studies	Biopharma for identifying highly expressed targets	Plant & animal whole genome annotation	
Experimental goal	Isoform discovery and quantification of moderate-to-rare transcripts	Isoform discovery of high expressed transcripts	Comprehensive transcript annotation in a species	
Example study design	Disease vs. normal tissues with multiple replicates  Disease vs. normal tissues with multiple replicates		Plant or animal with multiple tissue types	
Target depth of coverage per sample	10 M reads per sample	5 M reads per sample	≤5 M reads per tissue (of same species)	
Cample multipleving1	Sequel II/IIe system: Up to 2 samples per SMRT Cell 8M (2-plex)	Sequel II/IIe system: Up to 3 samples per SMRT Cell 8M (3-plex)	Sequel II/IIe system: Up to 3 tissue types per SMRT Cell 8M (3-plex)	
Sample multiplexing <sup>1</sup>	Revio system: Up to 4 samples per Revio SMRT Cell (4-plex)	Revio system: Up to 8 samples per Revio SMRT Cell (8-plex)	Revio system: Up to 8 tissue types per Revio SMRT Cell (8-plex) <sup>2</sup>	
SMRT Link data analysis workflows	Read Segmentation and Iso-Seq analysis application with option to "pool reads and cluster together" to get a master isoform classification file with per-sample full-length read counts			



### SMRT Link Read Segmentation and Iso-Seq analysis application setup

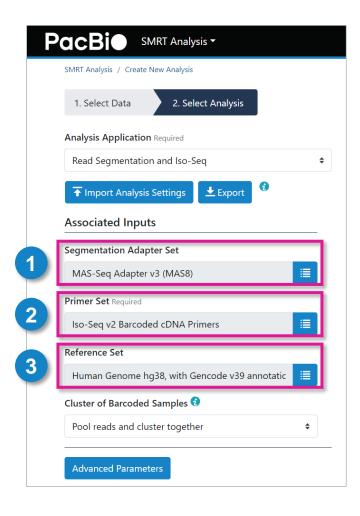
Specify Read Segmentation and Iso-Seq analysis application in SMRT Link





### SMRT Link Read Segmentation and Iso-Seq analysis application setup (cont.)

Specify Read Segmentation and Single-Cell Iso-Seq analysis application required associated inputs



#### Segmentation Adapter Set (Default = MAS-Seq Adapter v3 (MAS8))

 Specify a FASTA file, provided by PacBio, containing segmentation adapters. If you need a custom segmentation adapter set, click Advanced Parameters and use a custom FASTA file formatted as described in the <u>SMRT Link User Guide</u>

#### 2. Primer Set (Required) (Default = Iso-Seq v2 Barcoded cDNA Primers)

- Specify a primer sequence file in FASTA format to identify cDNA primers for removal. The primer sequence includes the 5' and 3' cDNA primers
- Primer IDs must be specified using the suffix \_5p to indicate 5' cDNA primers and the suffix \_3p to indicate 3' cDNA primers. The 3' cDNA primer should not include the Ts and is written in reverse complement. (See the <a href="SMRT Link User Guide">SMRT Link User Guide</a> for example Iso-Seq v2 Barcoded cDNA Primer IDs and sequences)
- Each primer sequence must be unique

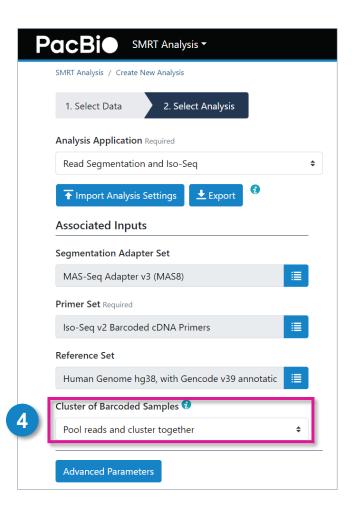
#### 3. Reference Set (Required)

- Specify one of two default reference genome and annotation sets to align high quality isoforms to, and to collapse isoforms mapped to the same genomic loci. The default sets are Human\_hg38\_Gencode\_v39 and Mouse\_mm39\_Gencode\_vM28 annotations
- Alternatively, choose other reference genomes (but not with annotations) that were custom-uploaded to SMRT Link
- The Reference Set can be left blank. If blank, the workflow will stop after the isoform clustering step (isoseq cluster)



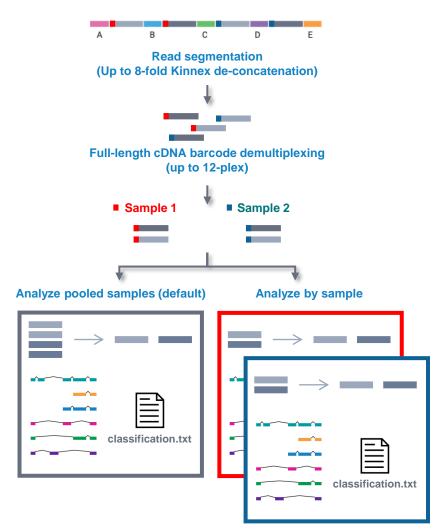
### SMRT Link Read Segmentation and Iso-Seq analysis application setup (cont.)

Specify Read Segmentation and Single-Cell Iso-Seq analysis application required associated inputs



#### 4. Cluster of Barcoded Samples

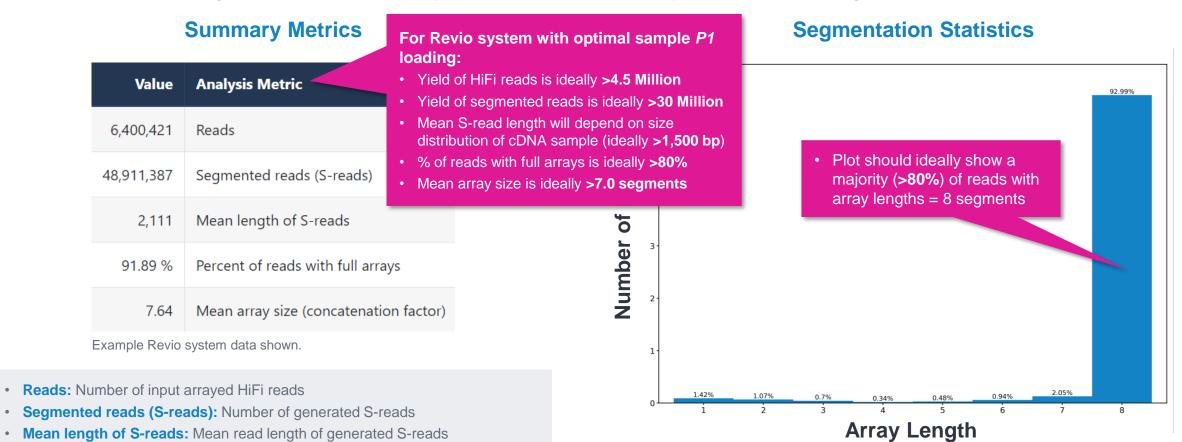
- This option specifies barcoded samples that were barcoded at the cDNA level, where the (barcoded) cDNA primers are specified in the Primer Set option. This option does not address libraries that were barcoded using barcoded adapters
- Specify whether all FLNC reads will be pooled for clustering, then demultiplexed based on pooled result. Note: This setting does not apply to non-barcoded samples
- Specify Pool reads and cluster together if barcoded samples are from the same species, but different tissues, or samples of the same genes but different individuals. The samples are clustered with all barcodes pooled
- Specify Cluster reads separately if barcoded samples are from different species. The samples are clustered separately by barcode
- In either case, the samples on the results page are automatically named BioSample\_1 through BioSample N





### Example SMRT Link Read Segmentation data utility processing results for Kinnex full-length RNA library prepared with human UHRR sample

SMRT Link Read Segmentation data utility job report – Summary Metrics and Segmentation Statistics



Histogram distribution of the number of S-reads per HiFi read. (Example Revio system data shown.)



Percent of reads with full arrays: Percentage of input HiFi reads containing all

adapter sequences in the order listed in the segmentation adapter FASTA file

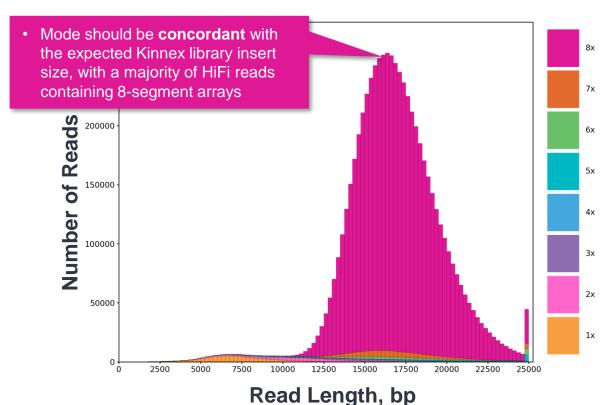
• Mean array size: Mean number of fragments (or S-reads) found in input reads

<sup>&</sup>lt;sup>1</sup> HiFi read lengths, reads/data per SMRT Cell and other sequencing performance results can vary depending on DNA sample quality, insert size, *P1* loading performance & movie time. Note: Refer to **SMRT Link MAS-Seq troubleshooting guide** (102-994-400) for example performance metrics typically achievable with Kinnex libraries under optimal *P1* loading conditions. For Sequel IIe systems, we recommend aiming for ~50 – 70% *P1* loading.

### Example SMRT Link Read Segmentation data utility processing results<sup>1</sup> for Kinnex full-length RNA library prepared with human UHRR sample (cont.)

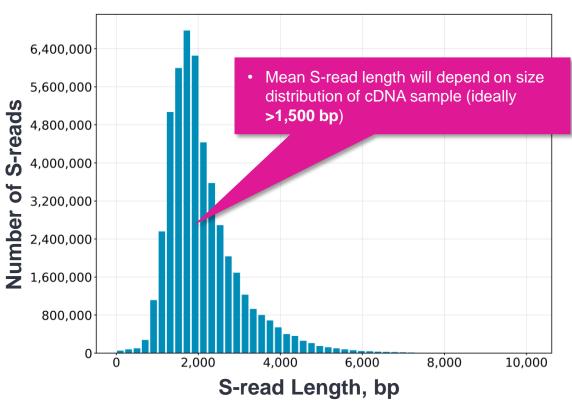
SMRT Link Read Segmentation data utility job report – Length of Reads and S-read Length Distribution

### **Length of Reads**



Histogram distribution of the number of HiFi reads by read length, in base pairs. (Example Revio system data shown.)

### **S-read Length Distribution**



Histogram distribution of the number of S-reads by HiFi read length, in base pairs. (Example Revio system data shown.)



<sup>1</sup> HiFi read lengths, reads/data per SMRT Cell and other sequencing performance results can vary depending on DNA sample quality, insert size, *P1* loading performance & movie time. Note: Refer to **SMRT Link MAS-Seq troubleshooting guide** (102-994-400) for example performance metrics typically achievable with Kinnex libraries under optimal *P1* loading conditions. For Sequel IIe systems, we recommend aiming for ~50 – 70% *P1* loading.

SMRT Link Iso-Seq analysis job report – Read Classification statistics

### **Summary Metrics**

Value	Analysis Metric
46,396,697	Reads
44,875,144	Reads with 5' and 3' Primers
44,784,675	Non-Concatamer Reads with 5' and 3' Primers
44,739,994	Non-Concatamer Reads with 5' and 3' Primers and Poly-A Tail (FLNC Reads)
1,981	Mean Length of FLNC Reads
12	Unique Primers
3,739,595	Mean Reads per Primer
4,379,484	Max. Reads per Primer
3,143,785	Min. Reads per Primer
1,521,553	Reads without Primers
96.22%	Percent Bases in Reads with Primers
96.72%	Percent Reads with Primers

Example Revio system data shown.

- Reads: Total number of CCS reads
- Reads with 5' and 3' Primers: Number of CCS reads with 5' and 3' cDNA primers detected
- Non-Concatemer Reads with 5' and 3' Primers: Number of nonconcatemer CCS reads with 5' and 3' primers detected
- Non-Concatemer Reads with 5' and 3' Primers and Poly-A Tail (FLNC Reads): Number of nonconcatemer CCS reads with 5' and 3' primers and polyA tails detected. This is usually the number for full-length, nonconcatemer (FLNC) reads, unless polyA tails are not present in the sample
- Mean Length of FLNC Reads: Mean length of the non-concatemer CCS reads with 5' and 3' primers and polyA tails detected
- Unique Primers: Number of unique primers in the sequence
- Mean Reads per Primer: Mean number of CCS reads per primer
- Max. Reads per Primer: Maximum number of CCS reads per primer
- Min. Reads per Primer: Minimum number of CCS reads per primer
- Reads without Primers: Number of CCS reads without a primer
- Percent Bases in Reads with Primers: Percentage of bases in CCS reads in the sequence data that contain primers
- Percent Reads with Primers: Percentage of CCS reads in the sequence data that contain primers



SMRT Link Iso-Seq analysis job report – Read Classification statistics

#### **Primer Data**

Bio Sample Name JT	Primer Name IT	CCS Pands IT	Mean Primer Quality JT	Pands with 5' and 2' Primare IT	Non-Concatamer Reads with 5' and 3' Primers 1	Non-Concatamer Reads wit
bio Sample Name 4	Trillier Name +	CC3 Reaus 4	ivicali Fillier Quality 4	Reads With 5 and 5 Filliers 4	Non-Concatanier Reads With 3 and 3 Timers 4	Non-Concatallier Reads with
BioSample_1	lsoSeqX_bc01_5plsoSeqX_3p	3,648,886	99.6	3,648,886	3,641,475	3,638,289
BioSample_2	IsoSeqX_bc02_5pIsoSeqX_3p	3,826,584	99.7	3,826,584	3,818,994	3,815,554
BioSample_3	IsoSeqX_bc03_5pIsoSeqX_3p	4,296,994	99.7	4,296,994	4,288,437	4,284,413
BioSample_4	IsoSeqX_bc04_5pIsoSeqX_3p	3,588,147	99.7	3,588,147	3,580,832	3,577,622
BioSample_5	IsoSeqX_bc05_5pIsoSeqX_3p	4,543,731	99.7	4,543,731	4,534,661	4,530,441
BioSample_6	IsoSeqX_bc06_5pIsoSeqX_3p	4,606,161	99.7	4,606,161	4,596,642	4,592,382
BioSample_7	IsoSeqX_bc07_5pIsoSeqX_3p	4,009,556	99.7	4,009,556	4,001,451	3,997,812
BioSample_8	lsoSeqX_bc08_5plsoSeqX_3p	3,651,440	99.7	3,651,440	3,644,176	3,641,097
BioSample_9	lsoSeqX_bc09_5plsoSeqX_3p	4,390,535	99.7	4,390,535	4,381,510	4,377,425
BioSample_10	lsoSeqX_bc10_5plsoSeqX_3p	4,049,234	99.7	4,049,234	4,041,043	4,037,382
BioSample_11	lsoSeqX_bc11_5plsoSeqX_3p	3,323,541	99.7	3,323,541	3,316,702	3,313,724
BioSample_12	lsoSeqX_bc12_5plsoSeqX_3p	3,372,840	99.7	3,372,840	3,365,834	3,362,745
Bio Sample 4	No Primer	1,603,738	0.0	0	0	0

Example Revio system data shown.

- Bio Sample Name: Name of the biological sample associated with the primer
- Primer Name: A string containing the pair of primer indices associated with this biological sample
- CCS Reads: Number of CCS reads associated with the primer
- Mean Primer Quality: Mean primer quality associated with the primer
- Reads with 5' and 3' Primers: Number of CCS reads with 5' and 3' cDNA primers detected
- Non-Concatemer Reads with 5' and 3' Primers: Number of non-concatemer CCS reads with 5' and 3' primers detected
- Non-Concatemer Reads with 5' and 3' Primers and Poly-A Tail: Number of non-concatemer CCS reads with 5' and 3' primers and polyA tails detected. This is usually the number for full-length, non-concatemer (FLNC) reads, unless polyA tails are not present in the sample.



SMRT Link Iso-Seq analysis job report – Read Classification statistics

Mean Number of Reads

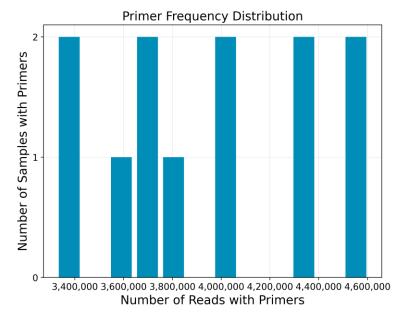
12

10

Example Revio system data shown.

# Number Of Reads Per Primer 4,650,000 4,350,000 4,350,000 3,900,000

#### **Primer Read Statistics**



Mean Read Length Distribution

Supplies the service of the service

**Number Of Reads Per Primer:** Maps the number of reads per primer, sorted by primer ranking

Primer Rank Order

**Primer Frequency Distribution:** Maps the number of samples with primers by the number of reads with primers

**Mean Read Length Distribution:** Maps the read mean length against the number of samples with primers

증 3,750,000

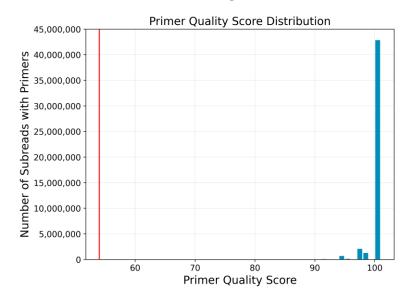
3,600,000

3,450,000

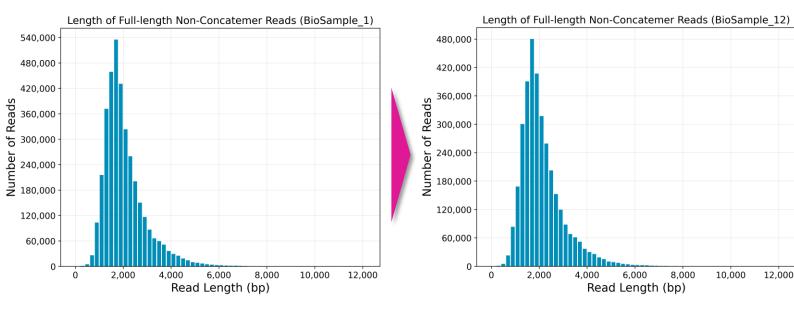
3,300,000

SMRT Link Iso-Seq analysis job report – Read Classification statistics

#### **Primer Quality Scores**



### **Length of Full-length Non-Concatemer Reads**



**Primer Quality Score Distribution:** Histogram of primer scores

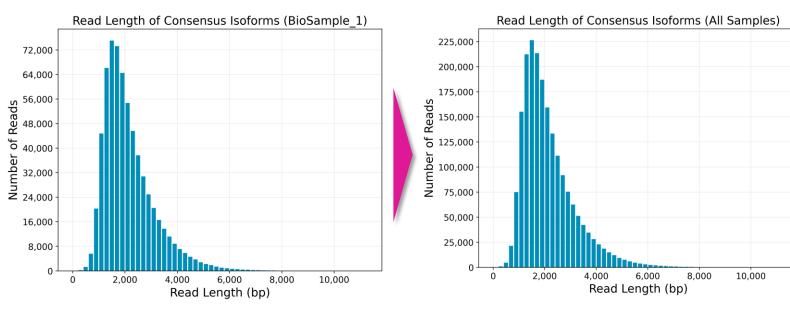
**Length of Full-Length Non-Concatemer Reads:** Per-sample histograms of the read length distribution of non-concatemer CCS reads with 5' and 3' primers and polyA tails detected

SMRT Link Iso-Seq analysis job report – Transcript Clustering statistics

### **Summary Metrics**

Sample Name ↓↑	Number of High-Quality Isoforms ↓↑
BioSample_1	650,495
BioSample_2	666,648
BioSample_3	712,671
BioSample_11	617,944
BioSample_12	621,931
All Samples	2,001,226

### **Length of Consensus Isoforms**



- Sample Name: Sample name for which the following metrics apply
- Number of High-Quality Isoforms: Number of consensus isoforms that have an estimated accuracy above the specified threshold
- Length of Consensus Isoforms: Per-sample histograms of the consensus isoform lengths and the distribution of isoforms exceeding a read length cutoff. Also includes a single histogram plot for all samples.

SMRT Link Iso-Seq analysis job report – Transcript Mapping and Classification statistics

### **Summary Metrics (All samples)**

Sample Name ↓↑	Total Unique Genes ↓↑	Total Unique Genes, filtered ↓↑	Total Unique Isoforms ↓↑	Total Unique Isoforms, filtered $\mbox{$\mathfrak{I}$}$
BioSample_1	127,560	18,576	317,862	127,517
BioSample_2	131,678	18,833	325,744	129,550
BioSample_3	143,056	19,089	348,790	135,483
BioSample_4	125,866	18,529	315,326	126,923
BioSample_5	148,422	19,451	360,453	138,918
BioSample_6	149,269	19,495	361,022	138,742
BioSample_7	136,195	18,867	336,140	132,612
BioSample_8	128,393	18,608	317,667	126,641
BioSample_9	145,921	19,291	353,923	136,903
BioSample_10	137,496	18,999	338,295	133,063
BioSample_11	120,079	18,308	301,435	122,466
BioSample_12	120,440	18,355	303,142	123,454

Example Revio system data shown.

- Sample Name: Sample name for which the following metrics apply
- Total unique genes: The total number of unique genes across all cells.
- Total unique genes, filtered: The total number of unique genes, after filtering out reads based on the SQANTI transcript filtering criteria.
- Total unique isoforms: The total number of unique isoforms across all cells
- Total unique isoforms, filtered: The total number of unique isoforms across all cells, after filtering out reads based on the SQANTI transcript filtering criteria.



SMRT Link Iso-Seq analysis job report – Transcript Mapping and Classification statistics

### **Transcript Classification, filtered (All samples)**

Category ↓↑	Count ↓↑	CAGE Detected ↓↑	CAGE Detected, (%)	) ↓↑   polyA Detected ↓↑	polyA Detected, (%) ↓↑
FSM	171662	83485	48.63%	92562	53.92%
ISM	257444	32719	12.70%	158888	61.71%
NIC	149131	97650	65.47%	78854	52.87%
NNC	106396	67432	63.37%	57440	53.98%
Antisense	1501	324	21.58%	923	61.49%
Fusion	3446	2021	58.64%	1901	55.16%
More junctions	83	49	59.03%	55	66.26%
Genic intron	0	0	0.00%	0	0.00%
Genic genomic	1166	524	44.93%	704	60.37%
Intergenic	3073	285	9.27%	2441	79.43%

Example Revio system data shown. Note: Unfiltered transcript classification data are also displayed in the Iso-Seq analysis job report.

- Category: Transcript classification assigned by the classification and filtering tool pigeon, based on the SQANTI3 software
- Count: Number of transcripts, after filtering out reads based on the SQANTI filtering criteria, in a specific classification
- CAGE Detected: Number of transcripts, after filtering, where the transcription start site falls within 50 bp of an annotated CAGE peak site
- CAGE Detected (%): Percentage of transcripts, after filtering, where the transcription start site falls within 50 bp of an annotated CAGE peak site
- polyA Motif Detected: Number of transcripts, after filtering, where a known polyA motif is detected upstream of the transcription end site
- polyA Motif Detected (%): Percentage of transcripts, after filtering, where a known polyA motif is detected upstream of the transcription end site



Filter out reads based on the

SQANTI3 transcript filtering criteria

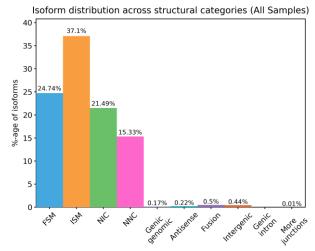
SMRT Link Iso-Seq analysis job report – Transcript Mapping and Classification statistics

### **Transcript Classification Plots**

# Isoform distribution across structural categories (All Samples) 27.88% 26.86% 27.88% 26.86%

Isoform length (kb)

#### **Transcript Classification Plots, Filtered**



### Structural categories by isoform lengths:

Isoform distributions across

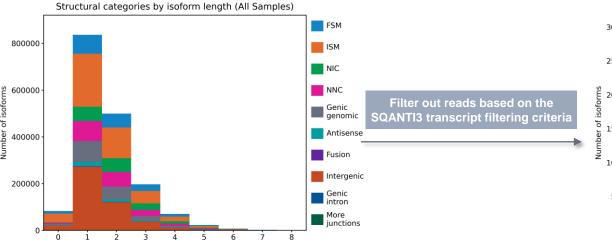
Distribution of the % of

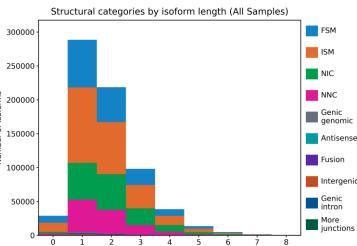
isoforms by structural

structural categories:

categories

 Histogram display of the number of isoforms by their length in kb and their structural category



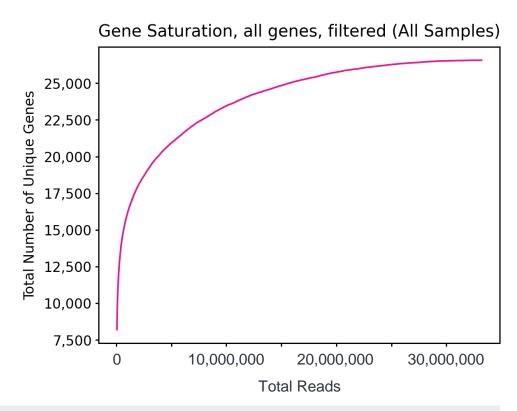


Isoform length (kb)

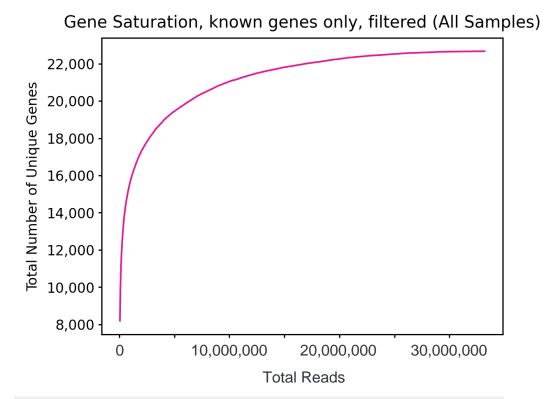


SMRT Link Iso-Seq analysis job report – Transcript Mapping and Classification statistics

#### **Gene Saturation**



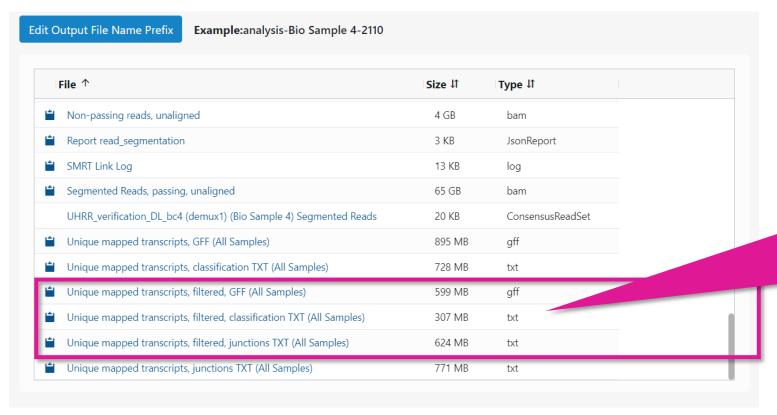
**Gene Saturation, all genes, filtered:** Saturation plot showing the level of gene saturation for all genes, after filtering out reads based on the SQANTI transcript filtering criteria



**Gene Saturation, known genes only, filtered:** Saturation plot showing the level of gene saturation, for unique known genes only (genes annotated in the reference annotation) per cell, after filtering out reads based on the SQANTI transcript filtering criteria

Example Revio system data shown.

#### File Downloads tab



Files shown in the File Downloads tab are available on the analysis results page. Additional files are also available on the SMRT Link server in the analysis output directory.

Refer to <u>SMRT Link user guide</u> for descriptions of downloadable output files

- These files are useful for visualizing isoform structures in Integrative Genomics Viewer (IGV) / UCSC genome browser and enable understanding of why an isoform is novel/known, etc.
  - GFF file containing unique mapped transcripts after filtering
  - Text file containing unique mapped transcript classifications against annotations, after filtering
  - Text file containing information about unique mapped transcript junctions, after filtering.

# Technical documentation & applications support resources

### Technical resources for Kinnex full-length RNA library preparation, sequencing & data analysis

### RNA sample preparation resources

 Note: The products below have not been tested or validated by PacBio but are listed here as examples of third-party kits used by other PacBio customers for isolating total RNA for PacBio RNA sequencing (Iso-Seq) applications

Kit type	Product name		
mRNA isolation	Ambion Poly(A) Purist MAG Kit [ Link ]		
	Qiagen RNeasy Plus Kits [ Link ]		
Total DNIA inclution	Sigma Spectrum Plant Total RNA Kit [ Link ]		
Total RNA isolation	iNtRON Easy Spin Total RNA [ Link ]		
	TRIzol Reagent can be used to isolate total RNA from tissues or cells, including lipid-rich and difficult samples [ Link ]		
RNA stabilization & storage	RNALater is an aqueous, nontoxic tissue storage reagent that rapidly permeates tissues to stabilize and protect cellular RNA [ Link ]		



### Technical resources for Kinnex full-length RNA library preparation, sequencing & data analysis (cont.)

### Kinnex full-length RNA library preparation literature & other resources

- Application note Kinnex full-length RNA kit for isoform sequencing (<u>102-326-591</u>)
- Procedure & checklist Preparing Kinnex libraries using Kinnex full-length RNA kit (103-238-700)
- Technical overview Kinnex kits for single-cell RNA, full-length RNA and 16S rRNA sequencing (103-343-700)
- Technical overview Kinnex library preparation using Kinnex full-length RNA kit (103-344-700)
- Video tutorial SMRT Link Sample Setup and Run Design setup procedure for Kinnex kits [ <u>Link</u> ]
- Whitepaper Bulk and single-cell isoform sequencing for human disease research (102-326-576)

### Data analysis resources

- Application note Bioinformatics tools for full length isoform sequencing (102-326-593)
- SMRT Link MAS-Seq troubleshooting guide (<u>102-994-400</u>)
- SMRT Link v13.0 software installation guide (<u>103-009-000</u>)
- SMRT Link v13.0 user guide (<u>103-008-800</u>)
- SMRT Tools v13.0 reference guide (<u>103-008-900</u>)
- Video tutorial Read Segmentation and Iso-Seq workflow in SMRT Link [ <u>Link</u> ]



### Technical resources for Kinnex full-length RNA library preparation, sequencing & data analysis (cont.)

#### **Publications**

- Schertzer, M.D. et al. (2023) Cas13d-mediated isoform-specific RNA knockdown with a unified computational and experimental toolbox.
   BioRxiv preprint [ <u>Link</u> ]
- Al'Khafaji, A.M. et al. (2023) High-throughput RNA isoform sequencing using programmable cDNA concatenation. Nature biotechnology. [
   <u>Link</u>]

#### Webinars

- PacBio Iso-Seq social club webinar (2022) Introduction to Iso-Seq method [ <u>Link</u> ]
- PacBio Iso-Seq social club webinar (2022) SQANTI3 for isoform classification and annotation [ Link ]
- PacBio Iso-Seq social club webinar (2022) TappAS for isoform differential expression analysis [ <u>Link</u> ]

### Example PacBio data sets

Application	Dataset	Data type	PacBio system
Kinnex full-length RNA sequencing	Homo sapiens – universal human reference RNA (UHRR) [ <u>Link</u> ]	HiFi long read	Sequel II & Revio systems
	Homo sapiens – HG002 [ <u>Link</u> ]	HiFi long read	Revio system



## PacBio

www.pacb.com

Research use only. Not for use in diagnostic procedures. © 2024 Pacific Biosciences of California, Inc. ("PacBio"). All rights reserved. Information in this document is subject to change without notice. PacBio assumes no responsibility for any errors or omissions in this document. Certain notices, terms, conditions and/or use restrictions may pertain to your use of PacBio products and/or third-party products. Refer to the applicable PacBio terms and conditions of sale and to the applicable license terms at pacb.com/license. Pacific Biosciences, the PacBio logo, PacBio, Circulomics, Omniome, SMRT, SMRTbell, Iso-Seq, Sequel, Nanobind, SBB, Revio, Onso, Apton, and Kinnex are trademarks of PacBio.