

## Introduction

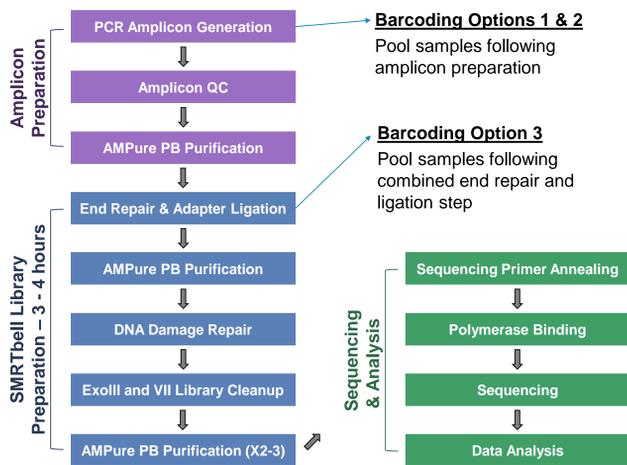
Targeted sequencing employing PCR amplification is a fundamental approach to studying human genetic disease. PacBio's Sequel System and supporting products provide an end-to-end solution for amplicon sequencing, offering better performance to Sanger technology in accuracy, read length, throughput, and breadth of informative data.

Sample multiplexing is supported with three barcoding options providing the flexibility to incorporate unique sample identifiers during target amplification or library preparation. Multiplexing is key to realizing the full capacity of the 1 million individual reactions per Sequel SMRT Cell.

Two analysis workflows, which can generate high-accuracy results, support a wide range of amplicon sizes. The Circular Consensus Sequencing workflow results in high accuracy through intra-molecular consensus generation, while high accuracy for the Long Amplicon Analysis workflow is achieved by clustering of individual long reads from multiple reactions.

Here we present workflows and results for single-molecule sequencing of amplicons for human genetic analysis.

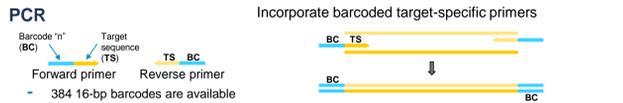
## Amplicon Library Preparation for Multiplex SMRT Sequencing



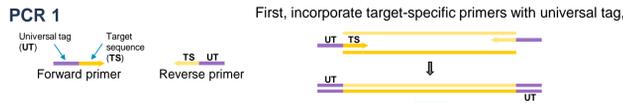
- Non-specific amplicons can be removed by AMPure PB purification or gel purification after amplicon generation or BluePippin or SageELF size selection after library preparation.

## Barcoding Options for Sample Multiplexing

### 1. Barcoded Primers

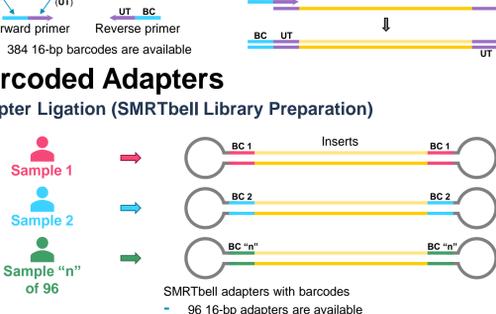


### 2. Barcoded Universal Primers



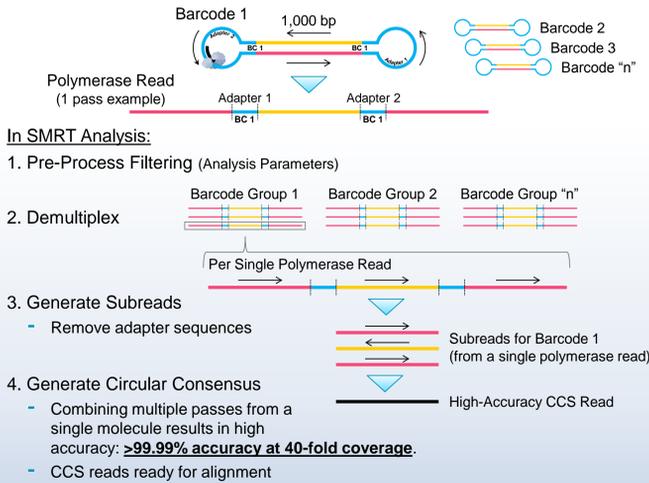
### 3. Barcoded Adapters

Adapter Ligation (SMRTbell Library Preparation)

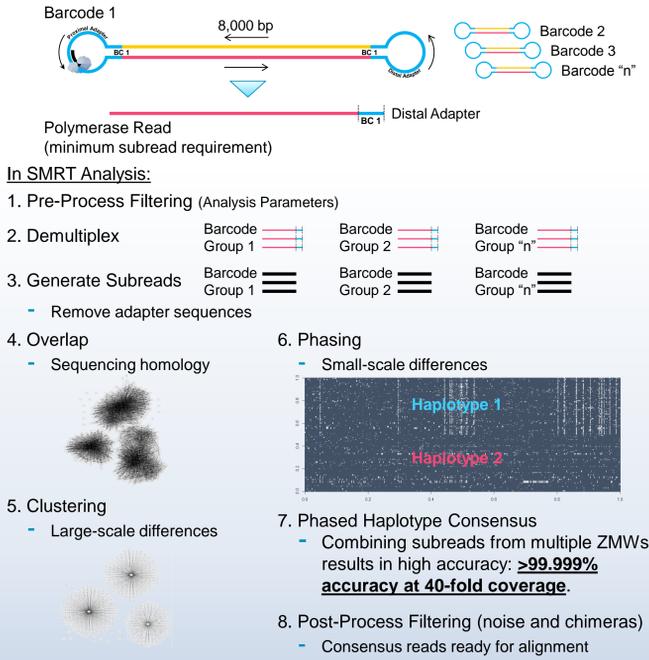


## Sequencing Analysis Workflows

### Circular Consensus Sequencing (CCS) Analysis



### Long Amplicon Analysis (LAA)

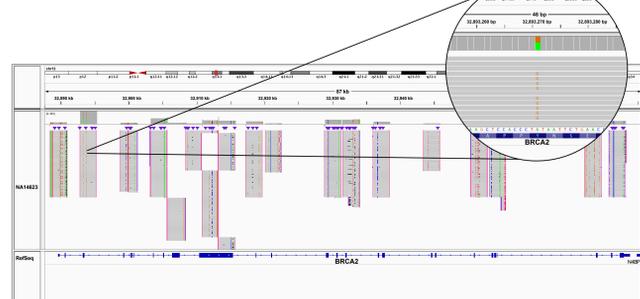


## Results – Mutation Detection in BRCA1/2 with CCS

The following results were generated from sequencing a 12-sample multiplex SMRTbell library on a Sequel System with each sample consisting of a 35-plex PCR covering BRCA1 and BRCA2 genes. Amplicons ranged in size from 2,300 to 2,800 bp, covering approximately 85 kb for each gene.

BC01 DNA	Gene	Known Mutation	
		Variant	Observed AF
NA14623	BRCA2	TYR42CYS	46.7%
NA14624	BRCA2	5946delCT	49.1%
NA14626	BRCA2	LYS3326TER	54.2%
NA13705	BRCA1	4-BP DEL, FS1252TER	45.4%
NA13715	BRCA1	1-BP INS, 5382C	42.9%
NA14090	BRCA1	2-BP DEL, 185AG	46.2%
NA14094	BRCA1	40-BP DEL, FS397TER	55.6%
NA14638	BRCA1	IVS5-11T>G	49.1%
NA14634	BRCA1	4-BP DEL, FS1364TER	51.3%
NA14636	BRCA1	5677insA	52.9%
NA14637	BRCA1	ARG1443TER	47.8%
NA14170	BRCA2	1-BP DEL, 6174T, FS	44.0%

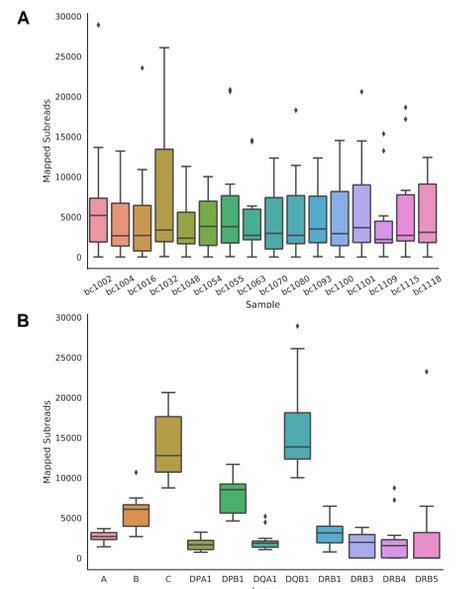
**Table 1. Mutations observed in samples with validated germline variants.** The expected variant was detected for each sample at the expected allele frequency indicating heterozygosity. "BC01 DNA" samples belong to a Coriell BRCA1/2 breast cancer mutation panel.



**Figure 1. An IGV alignment for NA14623 highlighting detection of TYR42CYS in BRCA2.** Eighteen amplicons span the BRCA2 region depicted above.

## Results – HLA Typing with LAA

The following results were generated from sequencing a 16-sample multiplex SMRTbell library on a Sequel System with each sample consisting of an 11-plex PCR covering HLA-A, -B, -C, -DPA1, -DPB1, -DQA1, -DQB1, -DRB1, and -DRB3/4/5 genes. Amplicons ranged in size from 3,500 to 6,500 bp.



**Figure 2. The number of mapped subreads compared across all 16 samples (A) and 11 loci (B).**

Locus	A	B	C	DPA1	DPB1	DQA1	DQB1	DRB1	DRB3	DRB4	DRB5
bc1002	100	100	100	100	100	100	100	100	100	100	100
bc1004	100	100	100	100	100	100	100	100	100	100	100
bc1016	100	100	100	100	100	100	100	100	100	100	100
bc1048	100	100	100	100	100	100	100	100	100	100	100
bc1054	100	100	100	100	100	100	100	100	100	100	100
bc1055	100	100	100	100	100	100	100	100	100	100	100
bc1063	100	100	100	100	100	100	100	100	100	100	100
bc1070	100	100	100	100	100	100	100	100	100	100	100
bc1080	100	100	100	100	100	100	100	100	100	100	100
bc1093	100	100	100	100	100	100	100	100	100	100	100
bc1100	100	100	100	100	100	100	100	100	100	100	100
bc1101	100	100	100	100	100	100	100	100	100	100	100
bc1109	100	100	100	100	100	100	100	100	100	100	100
bc1115	100	100	100	100	100	100	100	100	100	100	100
bc1118	100	100	100	100	100	100	100	100	100	100	100

**Table 2. Percent identity of cDNA sequence to reference.** Seven results highlighted in purple indicate previously unknown alleles. Note that sample bc1032 is homozygous across all loci.

Locus	A	B	C	DPA1	DPB1	DQA1	DQB1	DRB1	DRB3	DRB4	DRB5
bc1016	0	02:01:01:01	15:11:01	03:03:01	01:03:01:03	41:01:01	03:03:01	03:03:02:02	09:01:02	-	01:03:01
bc1032	1	02:06:01	39:01:03	07:02:01:01	01:03:01	03:01:01	03:02	04:01:01	04:05:01	-	01:03:02
bc1032	0	02:04	51:01:01:01	15:02:01:01	01:03:01:05	04:02:01:02	05:05:01	03:01:01	16:02:01	-	-
bc1032	1	-	-	-	-	-	-	-	-	-	02:05

**Table 3. Example HLA typing results for 2 of 16 samples.** Only one allele type was found for sample bc1032 at all loci.

## Conclusions

- Targeted amplicon sequencing is fully supported on the Sequel System.
- Three barcoding options enable multiplexing of samples, allowing for efficient use of SMRT Cells.
- Two analysis workflows, CCS and LAA, support amplicon sizes from 250 bp to >10 kb, producing high-accuracy results: >99.99% accuracy for CCS and >99.999% accuracy for LAA, both at 40-fold coverage.

## Resources / Acknowledgements

- Targeted Sequencing
- PacBio Webpage: [www.pacb.com/applications/targeted-sequencing/](http://www.pacb.com/applications/targeted-sequencing/)
- Barcoding
- Product Note, Barcoding Solutions: [Multiplexing Amplicons Up To 10 kb](#)
  - Document: [SMRT Analysis Barcoding Overview](#)
- Circular Consensus Sequencing
- Tutorial: [Circular Consensus Sequence analysis application](#)
- Long Amplicon Analysis
- Tutorial: [Long Amplicon Analysis application](#)

### Acknowledgements

BRCA primers provided by Swift Biosciences. The authors would like to thank everyone who helped generate poster data.