CAPTURE YOUR REGIONS OF INTEREST
IN HIGH RESOLUTION

PACBIO®

The PacBio® Systems combine single-molecule resolution, long reads, and uniform coverage, to **fully characterize genetic complexity** — including structural variants, indels, copy number variants, micro-satellites, somatic variants, haplotypes, and phasing. Move beyond simply cataloging SNPs and **target all types of variation across relevant genomic regions**, including low complexity regions like repeat expansions, promoters, and flanking regions of transposable elements.

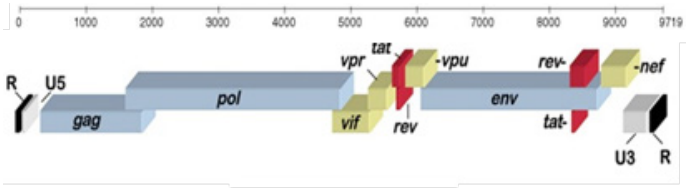
- Phase variants with allele specificity
- Sequence through low-complexity and highly repetitive regions
- Deconvolute haplotypes of complex populations with linear quantitation and minor variant detection below 0.1%
- Resolve structural variation breakpoints
- Accurately determine heterozygosity through *de novo* consensus generation
- Validate sequence variants cost-effectively and with high accuracy

RESOLUTION OF REPEAT EXPANSIONS

1. SMRT® Sequencing reads of the Fragile X region containing (CGG)n trinucleotide repeats.[1] An AGG repeat interruption (boxed) is unambiguously recognized in all reads.

2. A full Fragile X mutation allele of 750 CGG repeats.

SINGLE-BASE RESOLUTION OF COMPLEX MIXTURES OF FULL-LENGTH GENOMES



0 'AAAT+AAAAAAG+GT+GCA+ATTTACC+ACCC+' 25.2%

1 'AATT+AAAAAAG+AT+GGA+ATTTACC+ACCC+' 21.6%

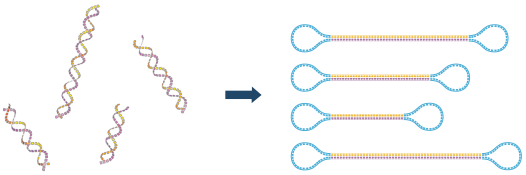
2 'AAAT+AAAAAAG+AT+GGA+ATTTACC+ACCC+' 21.0%

3 'AAAT+AGAAAAG+AT+GGA+ATTTACC+ACCC+' 19.6%

4 'AAAT+AAAAAAG+AT+GGA+ATTTTACC+ACCC+' 12.6%

Estimated haplotypes and relative abundances of five nearly identical full-length HIV-1 genomes (9.6 kb) from SMRT® Sequencing of a synthetic mixture of HIV-1 clones differing by only 1–2 bases.[2]

FROM TARGETED REGIONS TO ACCURATE VARIANTS AND HAPLOTYPES



Library Preparation

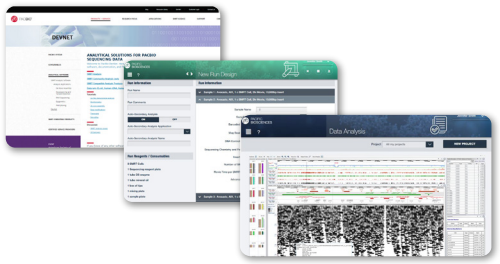
- Prepare SMRTbell™ libraries for a variety of insert sizes (250 bp to 40 kb)
- Utilize standard target-enrichment methods
- Multiplex 384 barcoded samples
- Enrich for longest inserts with size-selection options
- Support available for library automation

PacBio RS II System

Sequel System

SMRT® Sequencing with PacBio Systems

- Take advantage of the Sequel™ System to reduce project costs and generate 7X more reads compared with the PacBio RS II
- Achieve ~10 kb average read lengths, with some reads as long as 60 kb
- Adjust run times (0.5 to 6 hours) to maximize sample throughput and turn-around time
- Attain consensus accuracies > 99.999% by avoiding mapping and systematic errors
- Produce high intra-molecular accuracies with multiple observations of single circularized templates for complex population resolution



Data analysis with SMRT Analysis or PacBio DevNet

- Long Amplicon Analysis: generate reference-agnostic phased consensus assemblies from heterogeneous pools of amplicons
- Minor Variant Analysis: detect and quantify SNPs in a heterogeneous data set against a reference
- ClusterConsensus: reference-free deconvolution of genomes in a complex mixture
- HGAP: create high-quality *de novo* assemblies of plasmids, BACs or fosmids

KEY REFERENCES

1. Loomis, Erick W. et al. (2013) Sequencing the unsequenceable: Expanded CGG-repeat alleles of the fragile X gene. *Genome Research.* 23(1) 121-128.
2. Brown, M. et al. (2015, February) Sequencing complex mixtures of HIV-1 genomes with single-base resolution. *Poster presented at 16th annual Advances in Genome Biology & Technology.* Marco Island, FL.
3. Burke, G. R. et al. (2015) Microplitis demolitor Bracovirus proviral loci and clustered replication genes exhibit distinct DNA amplification patterns during replication. *Journal of Virology.* 89(18), 9511-9523.
4. Carneiro, M. O. et al. (2012) Pacific Biosciences sequencing technology for genotyping and variation discovery in human data. *BMC Genomics.* 13(1), 375.
5. 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(19), 11253–11263.
6. Mayor, Neema P et al. (2015) HLA typing for the next generation. *PLoS One.* 10(5), e0127153.
7. McFarland, Karen N. et al. (2015) SMRT Sequencing of long tandem nucleotide repeats in SCA10 reveals unique insight of repeat expansion structure. *PLoS One.* 10(8), e0135906.
8. Orkunoglu-Suer, F. et al. (2015) Targeted single molecule sequencing methodology for ovarian hyperstimulation syndrome. *BMC Genomics.* 16, 264.
9. Smith, C. et al. (2012) Validation of FLT3-ITD as a therapeutic target in human acute myeloid leukemia. *Nature.* 485(7397), 260-263.
10. Wang, M. et al. (2014) The genome sequence of African rice (Oryza glaberrima) and evidence for independent domestication. *Nature Genetics.* 46(9), 982–988.
11. Wang, M. et al. (2015) PacBio-LITS: a large-insert targeted sequencing method for characterization of human disease-associated chromosomal structural variations. *BMC Genomics.* 16, 214.

PN: AP101-092515