PacBio RS II Sequencing System

FIND MEANING IN COMPLEXITY

- Genome finishing
- Epigenetics
- Haplotype phasing
- Repeat expansions
- Isoform Sequencing
- Minor variants
Extraordinary Read Lengths with the PacBio RS II

The PacBio® RS II sequencing system allows scientists to rapidly and cost effectively generate finished genome assemblies, reveal and understand epigenomes, and characterize genomic variation. It achieves the industry’s longest read lengths and highest consensus accuracy.

Generate Finished Genomes

The PacBio RS II finishes microbial genomes and improves assembly of larger organisms with multi-kilobase reads and unbiased coverage regardless of GC content. No amplification is required.

Range of Genome Sizes

Benefits

- Resolve mobile elements and structural-variation events
- Generate complete, accurate, and contiguous genome assemblies
- Annotate more genes

De Novo Assembly Methods

- Hierarchical Assembly
- Hybrid Assembly
- Scaffolding
- Gap Filling

Paper: Genome Biology: Reducing assembly complexity of microbial genomes with single-molecule sequencing
Paper: Nature Methods: Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data
Paper: PloS One: Mind the gap: Upgrading genomes with Pacific Biosciences RS long-read sequencing technology
Paper: Applied and Environmental Microbiology: Genome of the anaerobic fungus Orpinomyces sp. C1A reveals the unique evolutionary history of a remarkable plant biomass degrader

www.pacb.com/denovo
Discover the Epigenome

The PacBio® RS II detects DNA base modifications using the kinetics of the polymerization reaction during sequencing.

Methyltransferases bind specifically to DNA motifs in a genome and methylate bases. PacBio software locates modified sites and motifs.

Methylome of the German E. coli outbreak strain. The inner and outer red circles show the kinetic signals. The colored internal tracks show the different methylation motif distributions.

<table>
<thead>
<tr>
<th>Motif</th>
<th>Occurrence in Genome</th>
<th>Modified in Genome</th>
<th>% Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'-GATC-3' 3'-CTAG-5'</td>
<td>42,992</td>
<td>41,969</td>
<td>97.6%</td>
</tr>
<tr>
<td>5'-ACCGAC-3' 3'-TGCTG-5'</td>
<td>4,569</td>
<td>4,492</td>
<td>98.3%</td>
</tr>
<tr>
<td>5'-CCTGAC-3' 3'-GACTG-5'</td>
<td>2,746</td>
<td>2,678</td>
<td>97.5%</td>
</tr>
<tr>
<td>5'-CCACN8TGAY-3' 3'-GCTGACTR-5'</td>
<td>492</td>
<td>478</td>
<td>97.2% 98.4%</td>
</tr>
</tbody>
</table>

Methyltransferases bind specifically to DNA motifs in a genome and methylate bases. PacBio software locates modified sites and motifs.

Genome-wide detection of methylation for the German E. coli outbreak strain.

Paper: Current Opinion in Microbiology: Entering the era of bacterial epigenomics with SMRT DNA sequencing
Paper: Nature Biotechnology: Genome-wide mapping of methylated adenine residues in pathogenic Escherichia coli using single-molecule real-time sequencing
Paper: Nucleic Acids Research: The methylomes of six bacteria
Characterize Genomic Variation

The PacBio® RS II’s long single-molecule reads and unbiased coverage provide access to the entire genome to accurately characterize genetic complexity.

**Isoform Sequencing**

Full-length sequencing of intact transcripts eliminates the need for read assembly, identifies splice variants, and improves gene annotation, to better understand the expressed gene set.

Poster: ABRF 2013: Full-length cDNA sequencing on the PacBio RS

**Repeat Expansions**

Span extreme CGG repeats and AT-rich regions with minimal bias, over hundreds to thousands of bases.

Paper: Genome Research: Sequencing the unsequenceable: Expanded CGG-repeat alleles of the fragile X gene

**Structural Variation and Copy-Number Variants**

With multi-kilobase reads, anchor the ends of duplicated and inverted regions of the genome, allowing direct analysis of structural variation.

Paper: American Journal of Respiratory Cell and Molecular Biology: Genome Reference and Sequence Variation in the Large Repetitive Central Exon of Human MUC5AC
Full-Length 16S Sequencing

Improve taxonomic resolution on metagenomic samples by sequencing full-length rRNA genes (16S, 18S) or other common species markers such as rpoB.

Poster: ASM 2013: Analysis of Full-Length Metagenomic 16S Genes by SMRT® Sequencing

Compound Mutations and Haploype Phasing

Study linked mutations hundreds, even thousands, of bases apart.

Paper: Nature: Validation of FLT3-ITD as a therapeutic target in human acute myeloid leukemia

Minor Variants and Quasispecies

Single molecule sequencing simplifies the analysis of mixed populations of sequences. Exquisitely sensitive and specific.

Poster: CROI 2013: Sensitive detection of minor variants and viral haplotypes using SMRT sequencing

SNP Detection and Validation

Highly accurate SNP validation for any genomic region reduces false positives and false negatives.

Paper: BMC Genomics: Pacific Biosciences sequencing technology for genotyping and variation discovery in human data
PacBio® RS II Typical Results

The PacBio RS II sequencing chemistries provide read lengths in excess of 20 kb with high consensus accuracy. P4-C2 achieves 99.999% consensus accuracy, ideal for de novo assembly and targeted sequencing applications. P5-C3 generates more 20 kb reads, best for scaffolding and spanning structural rearrangements.

**P4-C2 Chemistry**

- Average: ~5.5 kb
- Maximum: >24 kb
- Top 5% of reads: >11 kb
- Half of data in reads: >8 kb
- Data per SMRT® Cell: ~275 Mb

**P5-C3 Chemistry**

- Average: ~8.5 kb
- Maximum: >30 kb
- Top 5% of reads: >18 kb
- Half of data in reads: >10 kb
- Data per SMRT® Cell: ~375 Mb

Based on data from a 20 kb size-selected E. coli library using a 180-minute movie. Each SMRT Cell yields ~50,000 reads.

---

**Template Preparation**

<table>
<thead>
<tr>
<th>Insert Size (bp)</th>
<th>Input DNA per prep (ng)</th>
</tr>
</thead>
<tbody>
<tr>
<td>250 – 500</td>
<td>250</td>
</tr>
<tr>
<td>1,000 – 2,000</td>
<td>500</td>
</tr>
<tr>
<td>5,000 – 10,000</td>
<td>1,000</td>
</tr>
<tr>
<td>15,000 – 20,000</td>
<td>(size-selected) 5,000</td>
</tr>
</tbody>
</table>

Each library prep typically supports >35 SMRT Cells.
5 µg input material yields a minimum of >5 SMRT Cells with a size-selected library.

---

**Accuracy**

Data generated with P4-C2 chemistry on PacBio® RS II; Resequencing analysis using SMRT Analysis v2.0.1.
The PacBio® RS II system, consumables and software provide a complete set of tools to quickly and easily perform cutting-edge molecular biology.

**Reagents**
- DNA Template Prep Kit
- DNA Polymerase Binding Kit
- MagBead Kit
- AMPure® PB beads

**Instrument Run**
- PacBio RS II with touch screen
- RS Remote for run design
- SMRT Cells
- DNA Sequencing Kit

**Data Analysis**
- SMRT Analysis
- SMRT Portal
- SMRT View

**Sequencing time**
30 to 180 min per SMRT Cell

From DNA to data in as few as 10 hours

**SMRT® Technology**

The PacBio RS II is based on novel Single-Molecule, Real-Time (SMRT) technology which enables the observation of DNA synthesis by a DNA polymerase in real time. Sequencing occurs on SMRT Cells, each containing thousands of Zero-Mode Waveguides (ZMWs) in which polymerases are immobilized. The ZMWs provide a window for watching the DNA polymerase as it performs sequencing by synthesis.
Operating Environment

**Instrument and environmental cabinet**
- Power requirements: 208 – 240 VAC. UPS recommended
- Operating temperature: 15 °C – 25 °C (59 °F – 77 °F) ± 2 °C per hour
- Humidity: 20% – 80%, noncondensing
- Ventilation: HVAC capacity of up to 22,720 BTU (6654 Watts)
- Nitrogen: 90 – 125 PSI (4,654 – 6,464 torr)
- WxDxH: 78.9 in x 30.3 in x 62.2 in (200.4 cm x 77.0 cm x 158.0 cm)
- Weight: 2,405 lb (1,091 kg)

**Blade Center**
- Includes integrated computation and storage for performing single molecule, real-time sequencing, basecalling and quality assessment.
- WxDxH: 24.1 in x 35.9 in x 26.2 in (61.3 cm x 91.3 cm x 66.5 cm)
- Weight: 300 lb (136 kg)